

Manuscript version: Author's Accepted Manuscript

The version presented in WRAP is the author's accepted manuscript and may differ from the published version or Version of Record.

Persistent WRAP URL:

<http://wrap.warwick.ac.uk/107457>

How to cite:

Please refer to published version for the most recent bibliographic citation information. If a published version is known of, the repository item page linked to above, will contain details on accessing it.

Copyright and reuse:

The Warwick Research Archive Portal (WRAP) makes this work by researchers of the University of Warwick available open access under the following conditions.

Copyright © and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable the material made available in WRAP has been checked for eligibility before being made available.

Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

Publisher's statement:

Please refer to the repository item page, publisher's statement section, for further information.

For more information, please contact the WRAP Team at: wrap@warwick.ac.uk.

MINOR ARCS, MEAN VALUES, AND RESTRICTION THEORY FOR EXPONENTIAL SUMS OVER SMOOTH NUMBERS

ADAM J HARPER

ABSTRACT. We investigate exponential sums over those numbers $\leq x$ all of whose prime factors are $\leq y$. We prove fairly good minor arc estimates, valid whenever $\log^3 x \leq y \leq x^{1/3}$. Then we prove sharp upper bounds for the p -th moment of (possibly weighted) sums, for any real $p > 2$ and $\log^{C(p)} x \leq y \leq x$. Our proof develops an argument of Bourgain, showing this can succeed without strong major arc information, and roughly speaking it would give sharp moment bounds and restriction estimates for any set sufficiently factorable relative to its density.

By combining our bounds with major arc estimates of Drappeau, we obtain an asymptotic for the number of solutions of $a + b = c$ in y -smooth integers less than x , whenever $\log^C x \leq y \leq x$. Previously this was only known assuming the Generalised Riemann Hypothesis. Combining them with transference machinery of Green, we prove Roth's theorem for subsets of the y -smooth numbers, whenever $\log^C x \leq y \leq x$. This provides a deterministic set, of size $\approx x^{1-c}$, inside which Roth's theorem holds.

1. INTRODUCTION

One of the key tools in analytic number theory is the use of exponential sums over arithmetically interesting sets \mathcal{S} , namely sums of the form

$$\sum_{\substack{n \leq x, \\ n \in \mathcal{S}}} e(n\theta),$$

where $\theta \in \mathbb{R}$ and $e(z) := e^{2\pi iz}$ is the complex exponential. Such sums serve as Fourier transforms of \mathcal{S} , and are particularly useful in additive problems since one can detect solutions of additive equations very neatly in terms of exponential sums.

Exponential sums over primes are crucial in proving the ternary Goldbach conjecture, and also in proving results like Roth's theorem in the primes, which asserts that any positive density subset of the primes must contain a non-trivial three term arithmetic progression. Similarly, when \mathcal{S} is the set of k -th powers such sums have been hugely studied in connection with Waring's problem. In this paper we investigate exponential sums over the set $\mathcal{S}(y)$ of y -smooth numbers, that is the set of numbers all of whose

Date: 2nd October 2015.

2010 Mathematics Subject Classification. 11L07, 11N25, 11D04.

Key words and phrases. Exponential sums, circle method, restriction theory, smooth numbers, xyz conjecture, Roth's theorem.

The author is supported by a research fellowship at Jesus College, Cambridge.

prime factors are less than or equal to y . Smooth numbers appear throughout analytic number theory, for example in proving some of the sharpest results on Waring's problem, and in analysing the performance of cryptographic algorithms. This is described in the survey articles of Granville [15] and Hildebrand and Tenenbaum [23].

Before giving a precise description of our results and their consequences, we try to indicate what is at stake by considering a particular additive problem, and the type of exponential sum information one would need to solve it.

Suppose we were seeking an asymptotic or lower bound for the number $N(x, y)$ of solutions to the equation $a + b = c$, with $a, b, c \leq x$ all y -smooth numbers. This is a much studied problem with connections to the *abc* conjecture, as we will describe later. As usual in the circle method, the orthogonality of additive characters $e(n\theta)$ lets us write

$$N(x, y) = \int_0^1 \left(\sum_{\substack{n \leq x, \\ n \in \mathcal{S}(y)}} e(n\theta) \right)^2 \overline{\left(\sum_{\substack{n \leq x, \\ n \in \mathcal{S}(y)}} e(n\theta) \right)} d\theta.$$

Let us also note that, heuristically, we expect to have $N(x, y) \asymp \Psi(x, y)^3/x$ on a very wide range of y , where $\Psi(x, y) := \sum_{\substack{n \leq x, \\ n \in \mathcal{S}(y)}} 1$. This is because there are $\Psi(x, y)^3$ unrestricted choices of a, b, c , and the condition that $a + b = c$ might be expected to hold with “probability” of order $1/x$ over all such choices, since a, b, c have size at most x .

Now the usual way to proceed is to define a subset $\mathfrak{M} \subseteq [0, 1]$ of *major arcs*, typically a union of small intervals around rational numbers a/q with small denominators. One chooses \mathfrak{M} such that one can obtain asymptotics for $\sum_{\substack{n \leq x, \\ n \in \mathcal{S}(y)}} e(n\theta)$ when $\theta \in \mathfrak{M}$. Moreover, one usually expects the largest values of the exponential sum to occur on the major arcs, and therefore that they will make the dominant contribution to the integral.

Let $\mathfrak{m} := [0, 1] \setminus \mathfrak{M}$ denote the complementary set of *minor arcs*. In view of the above discussion, one would usually proceed by writing

$$\begin{aligned} N(x, y) &= \int_{\mathfrak{M}} \left(\sum_{\substack{n \leq x, \\ n \in \mathcal{S}(y)}} e(n\theta) \right)^2 \overline{\left(\sum_{\substack{n \leq x, \\ n \in \mathcal{S}(y)}} e(n\theta) \right)} d\theta + O\left(\sup_{\theta \in \mathfrak{m}} \left| \sum_{\substack{n \leq x, \\ n \in \mathcal{S}(y)}} e(n\theta) \right| \int_0^1 \left| \sum_{\substack{n \leq x, \\ n \in \mathcal{S}(y)}} e(n\theta) \right|^2 d\theta \right) \\ &= \int_{\mathfrak{M}} \left(\sum_{\substack{n \leq x, \\ n \in \mathcal{S}(y)}} e(n\theta) \right)^2 \overline{\left(\sum_{\substack{n \leq x, \\ n \in \mathcal{S}(y)}} e(n\theta) \right)} d\theta + O\left(\sup_{\theta \in \mathfrak{m}} \left| \sum_{\substack{n \leq x, \\ n \in \mathcal{S}(y)}} e(n\theta) \right| \Psi(x, y) \right), \end{aligned}$$

where the second equality uses Parseval's identity to bound the mean square.

We now come to the crucial point. Parseval's identity, although simple and general, is also extremely powerful, since it shows that on average the exponential sum $\sum_{\substack{n \leq x, \\ n \in \mathcal{S}(y)}} e(n\theta)$ exhibits squareroot cancellation. Nevertheless, since we expect that

$N(x, y)$ (and therefore also the integral over the major arcs) should have size $\asymp \Psi(x, y)^3/x$, in order to obtain a non-trivial result we would need to know (somewhat more than)

$$\sup_{\theta \in \mathfrak{m}} \left| \sum_{\substack{n \leq x, \\ n \in \mathcal{S}(y)}} e(n\theta) \right| \ll \frac{\Psi(x, y)^2}{x}.$$

Here the trivial bound is $\Psi(x, y)$, so we ask for a saving of $\Psi(x, y)/x$. But if $y = \log^K x$, say, for any fixed $K \geq 1$, then it is known that $\Psi(x, \log^K x) = x^{1-1/K+o(1)}$ as $x \rightarrow \infty$, and so we need a *power saving* $x^{-1/K+o(1)}$ on the minor arcs. It is almost never possible to define a set of major arcs on which one can obtain asymptotics, and which is large enough that one has a power saving on the complementary minor arcs, except by assuming very strong conjectures like the Generalised Riemann Hypothesis. Note that if one were studying a ternary additive problem involving the primes, say, then one would only need to save a factor of about $1/\log x$ on the minor arcs, corresponding to the density of the primes. In this sense problems involving primes are much easier than problems involving sparse sets like very smooth numbers.

At this point, a circle method enthusiast might suggest various *pruning* manoeuvres to try to make progress. We will discuss this possibility later, but it has not yet made progress on the above problem, and our approach in this paper is different.

Note that, similarly as above, we can write

$$N(x, y) = \int_{\mathfrak{M}} \left(\sum_{\substack{n \leq x, \\ n \in \mathcal{S}(y)}} e(n\theta) \right)^2 \overline{\left(\sum_{\substack{n \leq x, \\ n \in \mathcal{S}(y)}} e(n\theta) \right)} d\theta + O \left(\sup_{\theta \in \mathfrak{m}} \left| \sum_{\substack{n \leq x, \\ n \in \mathcal{S}(y)}} e(n\theta) \right|^{0.9} \int_0^1 \left| \sum_{\substack{n \leq x, \\ n \in \mathcal{S}(y)}} e(n\theta) \right|^{2.1} d\theta \right),$$

say. Heuristically one might now expect that

$$\int_0^1 \left| \sum_{\substack{n \leq x, \\ n \in \mathcal{S}(y)}} e(n\theta) \right|^{2.1} d\theta \ll \max \left\{ \frac{\Psi(x, y)^{2.1}}{x}, \Psi(x, y)^{1.05} \right\},$$

since the first term reflects a contribution from major arcs, e.g. points at distance $\leq 1/x$ from 0, and the second term reflects squareroot cancellation. Provided $y \geq \log^K x$ for sufficiently large fixed K , the term $\Psi(x, y)^{2.1}/x$ will be the dominant one. Now *if* we could show that indeed $\int_0^1 \left| \sum_{\substack{n \leq x, \\ n \in \mathcal{S}(y)}} e(n\theta) \right|^{2.1} d\theta \ll \frac{\Psi(x, y)^{2.1}}{x}$, then in order for the minor arc contribution to be small we would only need to show that

$$\sup_{\theta \in \mathfrak{m}} \left| \sum_{\substack{n \leq x, \\ n \in \mathcal{S}(y)}} e(n\theta) \right|^{0.9} = o(\Psi(x, y)^{0.9}).$$

In other words, we would no longer need to obtain a power saving, but rather *any* saving at all would be sufficient.

In this paper we will prove sharp mean value estimates $\int_0^1 \left| \sum_{\substack{n \leq x, \\ n \in \mathcal{S}(y)}} e(n\theta) \right|^p d\theta \ll \frac{\Psi(x,y)^p}{x}$ when $p > 2$. In fact, we will prove a much more general *restriction* result for arbitrarily weighted mean values $\int_0^1 \left| \sum_{\substack{n \leq x, \\ n \in \mathcal{S}(y)}} a_n e(n\theta) \right|^p d\theta$. We will also prove some new pointwise minor arc bounds. These results have applications to ternary additive problems involving smooth numbers, and also to certain additive combinatorics questions.

We should emphasise that there is nothing new about trying to bound fractional moments of exponential sums. However, existing techniques would generally require an equivalent amount of major arc information as would make the standard circle method approach to ternary additive problems work directly. In our problems we do not have such information available. In contrast, we will obtain sharp mean value bounds in a quite general way, requiring very little major arc information.

1.1. Brief survey of previous results. Before explaining our new results, we indicate more precisely what was known about exponential sums over smooth numbers, with some comparisons to what is known about sums over primes.

Major arc estimates. If $\theta = a/q + \delta$, for some small $q \in \mathbb{N}$, some $(a, q) = 1$, and some small real δ , then θ is often described as lying on a *major arc*, and one can evaluate $\sum_{\substack{n \leq x, \\ n \in \mathcal{S}}} e(n\theta)$ quite precisely by splitting into arithmetic progressions to the small modulus q . In particular, if \mathcal{S} is “well distributed” in arithmetic progressions to modulus q then one can hope to obtain an asymptotic formula for $\sum_{\substack{n \leq x, \\ n \in \mathcal{S}}} e(n\theta)$.

For primes, one can show (see chapter 26 of Davenport [10]) that for any $A > 0$,

$$\sum_{n \leq x} \Lambda(n) e(n(a/q + \delta)) = \frac{\mu(q)}{\phi(q)} T(\delta) + O_A((1 + |\delta x|) x e^{-c\sqrt{\log x}}) \quad \forall q \leq \log^A x, (a, q) = 1,$$

where $\mu(q)$ denotes the Möbius function, $\phi(q)$ denotes Euler’s totient function, and $T(\delta) := \sum_{n \leq x} e(n\delta)$ is the sum of a geometric progression.

In the case of smooth numbers, Drappeau [12] recently showed that if $\log^{c_1} x \leq y \leq x$, and $q \leq y^{c_2}$, and $(a, q) = 1$, for certain absolute constants $c_1, c_2 > 0$, then

$$\sum_{\substack{n \leq x, \\ n \in \mathcal{S}(y)}} e(n(a/q + \delta)) = \sum_{\substack{n \leq x, \\ n \in \mathcal{S}(y)}} \frac{\mu(q/(q, n))}{\phi(q/(q, n))} e(n\delta) + \text{error terms},$$

and he gave useful estimates for the sum on the right hand side. We state these results carefully in §2.2, since we shall need to use them. Drappeau’s estimates extend the permitted range of y (and q and δ) as compared with previous results, for example those of La Bretèche [4], La Bretèche and Granville [6], and Fouvry and Tenenbaum [13].

Minor arc estimates. If θ isn't close enough to a rational with small denominator to be treated by major arc methods, then θ is usually described as lying on the *minor arcs*, and we usually settle for obtaining non-trivial upper bounds for $\sum_{\substack{n \leq x \\ n \in \mathcal{S}}} e(n\theta)$.

For primes, the most common minor arc estimates (see chapter 25 of Davenport [10]) are of the following rough form, which is non-trivial provided $\log^8 x \ll q \ll x/\log^8 x$:

$$\left| \sum_{n \leq x} \Lambda(n) e(n(a/q + \delta)) \right| \ll x \log^4 x \left(\frac{1}{\sqrt{q}} + \frac{1}{x^{1/5}} + \sqrt{\frac{q}{x}} \right) \quad \forall q \in \mathbb{N}, (a, q) = 1, |\delta| \leq \frac{1}{q^2}.$$

One can also prove stronger results with an explicit dependence on δ (e.g. Helfgott [20] exploits such results in his work on the ternary Goldbach conjecture, building on their use elsewhere in the circle method). All these ultimately depend on formatting $\sum_{n \leq x} \Lambda(n) e(n(a/q + \delta))$ into multiple sums, as pioneered by Vinogradov, and then using the Cauchy–Schwarz inequality and completing one sum. The bound decays like $1/\sqrt{q}$ (and, in the δ -dependent estimates, roughly like $1/\sqrt{q(1 + |\delta x|)}$), but we expect decay roughly like $1/q(1 + |\delta x|)$ on a wide range, as proved on a limited range of q and δ by the major arc estimates. This loss is ultimately due to the use of Cauchy–Schwarz, a key point we shall return to later.

For smooth numbers, the best results that are comparable to the prime number estimates seem to be the following bounds, in which $q \in \mathbb{N}$ and $(a, q) = 1$:

$$\left| \sum_{\substack{n \leq x, \\ n \in \mathcal{S}(y)}} e(n(\frac{a}{q} + \delta)) \right| \ll \begin{cases} x(1 + |\delta x|) \log^3 x \left(\frac{\sqrt{y}}{x^{1/4}} + \frac{1}{\sqrt{q}} + \sqrt{\frac{qy}{x}} \right) & \text{if } 3 \leq y \leq \sqrt{x}, \\ x \sqrt{\log x} \log y \left(\sqrt{\frac{y}{x}} + \frac{1}{\sqrt{q}} + \sqrt{\frac{q}{x}} + e^{-\sqrt{\log x}} \right) & \text{if } x, y \geq 2, |\delta| \leq \frac{1}{q^2}. \end{cases} \quad (1.1)$$

These are due to Fouvry and Tenenbaum [13] and to La Bretèche [4], respectively. They are a bit unsatisfactory unless y and q are fairly large, because the leading term is x rather than the trivial bound of $\Psi(x, y) := \sum_{\substack{n \leq x, \\ n \in \mathcal{S}(y)}} 1$, which may be *much* smaller.

Our first theorem, stated shortly, will be a minor arc estimate where the leading term is $\Psi(x, y)$, valid on the wide range $\log^3 x \leq y \leq x^{1/3}$, and with an explicit dependence on δ that further strengthens the bound on certain ranges.

Mean value theorems. For many sets \mathcal{S} we expect a mean value bound

$$\int_0^1 \left| \sum_{\substack{n \leq x, \\ n \in \mathcal{S}}} e(n\theta) \right|^p d\theta \ll_p \max \left\{ \frac{1}{x} \left(\sum_{\substack{n \leq x, \\ n \in \mathcal{S}}} 1 \right)^p, \left(\sum_{\substack{n \leq x, \\ n \in \mathcal{S}}} 1 \right)^{p/2} \right\}$$

for any real $p \geq 2$, since the first term reflects a contribution from major arcs, e.g. points at distance $\leq 1/x$ from 0, and the second term reflects squareroot cancellation.

Note that when $p = 2$ the second term is always the dominant one, but for larger p the first term may dominate.

For primes, it is fairly easy to show that indeed $\int_0^1 |\sum_{n \leq x} \Lambda(n) e(n\theta)|^p d\theta \ll_p x^{p-1}$ whenever $p > 2$, simply by adapting the argument we outlined when thinking about $N(x, y)$. For if \mathfrak{M} denotes a suitable set of major arcs (whose definition will depend on p), and $\mathfrak{m} := [0, 1] \setminus \mathfrak{M}$ the complementary set of minor arcs, then

$$\int_0^1 \left| \sum_{n \leq x} \Lambda(n) e(n\theta) \right|^p d\theta \ll \int_{\mathfrak{M}} \left| \sum_{n \leq x} \Lambda(n) e(n\theta) \right|^p + \sup_{\theta \in \mathfrak{m}} \left| \sum_{n \leq x} \Lambda(n) e(n\theta) \right|^{p-2} x \log x,$$

by Parseval's identity. One can evaluate the integral over \mathfrak{M} using pointwise asymptotics, and one has bounds for the supremum on the minor arcs that save at least a factor $1/\log^{1/(p-2)} x$. The crucial point here is that the primes are logarithmically far from being a dense set, producing the unwanted factor $\log x$, and one has major arc estimates on a wide enough range of q and δ to compensate for that logarithmic loss.

For smooth numbers, when $C\sqrt{\log x} \log \log x \leq \log y$ one could prove a sharp bound $\int_0^1 \left| \sum_{\substack{n \leq x, \\ n \in \mathcal{S}(y)}} e(n\theta) \right|^p d\theta \ll_p \Psi(x, y)^p / x$ for any $p > 2$, as for the primes. (See §2.1 for some explanation of this range.) But for smaller y the known major arc estimates are valid on too small a range of q to compensate for the very low density of the y -smooth numbers, even when accompanied by our new minor arc estimates. Thus it has been impossible to handle many additive problems involving smooth numbers when y is small.

Our second and main theorem will be a sharp upper bound for $\int_0^1 \left| \sum_{\substack{n \leq x, \\ n \in \mathcal{S}(y)}} e(n\theta) \right|^p d\theta$, that is valid whenever $p > 2$ and $\log^{C(p)} x \leq y \leq x$.

Restriction theory. Lastly we reach a class of results that may be less familiar to some analytic number theorists, namely *majorant* and *restriction* theorems. Some of this terminology is imported from Euclidean harmonic analysis, but all we mean is that we would like to show, for any complex numbers $|a_n| \leq 1$ and any $p > 2$, and for a given nice set \mathcal{S} , that

$$\int_0^1 \left| \sum_{\substack{n \leq x, \\ n \in \mathcal{S}}} a_n e(n\theta) \right|^p d\theta \ll_p \int_0^1 \left| \sum_{\substack{n \leq x, \\ n \in \mathcal{S}}} e(n\theta) \right|^p d\theta, \quad \text{or} \quad \int_0^1 \left| \sum_{\substack{n \leq x, \\ n \in \mathcal{S}}} a_n e(n\theta) \right|^p d\theta \ll_p \left(\sum_{\substack{n \leq x, \\ n \in \mathcal{S}}} 1 \right)^p / x.$$

This is another mean value problem, but the point is that we would like results that don't require the a_n to be “nice”, but hold simply because \mathcal{S} has good properties.

For primes, the restriction estimate

$$\int_0^1 \left| \sum_{p \leq x} a_p e(p\theta) \right|^q d\theta \ll_q \pi(x)^q / x \quad \forall q > 2$$

was first proved by Bourgain [3] (though expressed rather differently, and with an L^2 rather than an L^∞ condition on the coefficients a_p), and used to show that random subsets of the primes have certain nice harmonic analysis properties. We will discuss his proof later. These results gained great prominence when Green [16] gave another proof, and used the restriction estimate to prove Roth's theorem in the primes.

For smooth numbers, the author is not aware of any previous restriction or majorant results, though they could have been proved for large y (e.g. for $\log y \geq C\sqrt{\log x} \log \log x$) by direct use of Bourgain's [3] method. Green's [16] method crucially exploits the fact that primes can be majorised efficiently by sieve processes¹, and since that is well known to fail for smooth numbers (see e.g. §3.3 of Granville's survey [15]) his method doesn't seem applicable.

We shall prove a restriction theorem for the y -smooth numbers whenever $p > 2$ and $\log^{C(p)} x \leq y \leq x$. The key new point is that we can do this even though pointwise major and minor arc estimates are insufficient even to sharply bound the *unweighted* mean value. Indeed, we deduce our mean value theorem from our restriction theorem.

1.2. Statement of new results and applications. Firstly we shall prove a minor arc estimate that is sensitive to the density of the smooth numbers. This result involves a standard quantity $\alpha(x, y)$, whose definition and properties will be summarised in §2.1. For the moment, we only say very roughly that $\alpha(x, y)$ is “close to” 1 provided $(\log y)/\log \log x$ is large.

Theorem 1. *Let $\log^3 x \leq y \leq x^{1/3}$ be large. Suppose that $\theta = \frac{a}{q} + \delta$ for some $(a, q) = 1$ and some $\delta \in \mathbb{R}$. Provided that $q^2 y^3 (1 + |\delta x|)^2 \leq x/4$, we have the bound*

$$\left| \sum_{\substack{n \leq x, \\ n \in \mathcal{S}(y)}} e(n\theta) \right| \ll \frac{\Psi(x, y)}{\sqrt{q(1 + |\delta x|)}} (q(1 + |\delta x|))^{(3/2)(1 - \alpha(x, y))} u^{3/2} \log u \log x \sqrt{\log(2 + |\delta x|) \log(qy)},$$

where $\Psi(x, y) := \sum_{\substack{n \leq x, \\ n \in \mathcal{S}(y)}} 1$, where $u := (\log x)/\log y$, and where $\alpha(x, y)$ denotes the saddle-point corresponding to the y -smooth numbers less than x .

In particular, under the above conditions we have

$$\left| \sum_{\substack{n \leq x, \\ n \in \mathcal{S}(y)}} e(n\theta) \right| \ll \frac{\Psi(x, y)}{(q(1 + |\delta x|))^{1/2 - (3/2)(1 - \alpha(x, y))}} \log^{7/2} x.$$

For a reader who is less familiar with smooth numbers, we emphasise again that when y is small they are very sparse, for example $\Psi(x, \log^K x) = x^{1-1/K+o(1)}$ for any fixed

¹The characteristic function of the primes is decomposed into pieces such that, roughly speaking, earlier pieces are dense (so Parseval-type arguments are efficient) but not particularly Fourier uniform, whereas later pieces are sparser but only produce small exponential sums. This crucial decomposition comes from sieving with increasing levels.

$K \geq 1$, as $x \rightarrow \infty$. Thus it is important to have estimates in which the leading term is $\Psi(x, y)$ rather than x . The condition that $q^2 y^3 (1 + |\delta x|)^2 \leq x/4$, which arises naturally in the proof, is not really restrictive. That is because we can generally assume that $y \leq x^{1/100}$, and (after applying Dirichlet's approximation theorem to θ) that $q \leq x^{0.55}$ and $|\delta x| \leq x^{0.45}/q$, say. Then the only case we cannot handle with Theorem 1 is where $x^{0.48} \leq q \leq x^{0.55}$, for which we can use Fouvry and Tenenbaum's [13] minor arc estimate (stated earlier in (1.1)) and obtain a substantial saving because q is so large. Note also that, because our estimate improves as δ increases, we don't need to restrict to narrow arcs around a/q even when q is small, which is helpful in applications.

The proof of Theorem 1 isn't very difficult, since it works in the standard way of formatting the exponential sum into double sums, which is easy because the smooth numbers have good factorisation properties, and then using the Cauchy–Schwarz inequality. The only issue that arises is obtaining a reasonable upper bound for the quantity of y -smooth numbers in a segment of an arithmetic progression, on a wide range of y . By combining a method of Friedlander [14] with modern smooth number estimates we obtain upper bounds that aren't too bad (see §2.1, below), and by controlling the lengths of our double sums we can ensure the losses are acceptable.

Our second and main result is the following restriction theorem. Note that by choosing all the coefficients a_n to be 1 we obtain a mean value result for exponential sums over smooth numbers, which was itself unknown for small y .

Theorem 2. *There exists an absolute constant $C > 0$ such that the following is true.*

Let $p > 2$, suppose x is large enough in terms of p , and let $\log^{C \max\{1, 1/(p-2)\}} x \leq y \leq x$. Then for any complex numbers $(a_n)_{n \leq x}$ with absolute values at most 1, we have

$$\int_0^1 \left| \sum_{\substack{n \leq x, \\ n \in \mathcal{S}(y)}} a_n e(n\theta) \right|^p d\theta \ll_p \frac{\Psi(x, y)^p}{x}.$$

The lower bound condition $y \geq \log^{C \max\{1, 1/(p-2)\}} x$ is sharp up to the value of C , at least for p close to 2. For it is easy to show that $\int_0^1 \left| \sum_{\substack{n \leq x, \\ n \in \mathcal{S}(y)}} e(n\theta) \right|^p d\theta \gg_p \Psi(x, y)^{p/2}$ whenever $p \geq 2$, and since $\Psi(x, \log^{C/(p-2)} x) = x^{1-(p-2)/C+o(1)}$ we are close to the transition point where $\Psi(x, y)^{p/2} > \Psi(x, y)^p/x$.

To explain the proof of Theorem 2, we first explain Bourgain's [3] restriction argument for the primes. Bourgain [3] bounds the measure of the set of θ for which $\left| \sum_{p \leq x} a_p e(p\theta) \right| \geq \delta \pi(x)$, which will yield the restriction theorem if one can obtain good bounds for a sufficient range of δ . He makes a clever application of the Cauchy–Schwarz inequality, (described as “linearization”, and which the reader might recognise from Halász–Montgomery type arguments for Dirichlet polynomials), after which the

unknown coefficients a_p are removed and one can insert bounds for *unweighted* exponential sums over primes, with θ replaced by differences $\theta_r - \theta_s$. Since one has good major arc estimates for primes, one obtains bounds for the set of θ after finally exploiting the good spacing properties of the major arcs.

We will show that Bourgain’s argument can be run without inserting strong major arc estimates. In fact, we will show that his Cauchy–Schwarz “duality” step combines very naturally with the usual *proof* of minor arc estimates, that involves producing double sums. The key point is that, when proving pointwise bounds like Theorem 1, the use of the Cauchy–Schwarz inequality is essential but wasteful, losing a squareroot factor. In contrast, in Bourgain’s argument we already have a Cauchy–Schwarz step that is not wasteful, since it involves sets of θ rather than \mathcal{S} . It turns out that this one Cauchy–Schwarz step can suffice for everything, so that the double sums approach to minor arcs, which loses a squareroot factor when seeking pointwise bounds, becomes quite sharp when seeking mean value or restriction results. This approach is not at all restricted to smooth numbers, and though we don’t attempt a general formulation a rough indication of the requirement on $\mathcal{S} \subseteq [1, x]$ is the following: for any $1 \leq K \leq (\frac{x}{\#\mathcal{S}} \log x)^{100}$, one must be able to find some set $\mathcal{S}_K \subseteq [1, x/K]$ such that

$$\mathcal{S} \subseteq \{mn : m \in \mathcal{S}_K, n \leq x/m\}, \quad \text{and} \quad \sum_{m \in \mathcal{S}_K} \sum_{n \leq x/m} 1 \ll (\#\mathcal{S}) K^{0.01} \log^{1000} x.$$

(Here the numbers 100, 0.01, 1000, are indicative only, and will really depend on how close p is to 2 and on how much major arc information one has about \mathcal{S} .)

We actually use Theorem 1, combined with the Erdős–Turán inequality, in the proof of Theorem 2, to reduce to a situation where we only need look at differences $\theta_r - \theta_s$ that are “fairly close” to a rational a/q with “fairly small” denominator (a reduction that lets us set the lengths of double sums properly). We also use sharp major arc estimates, but only in the very restricted case where q has size $\log^{O(1)} x$, because our main argument loses a few logarithmic factors and we need a different approach if the other terms aren’t large enough to compensate for this.

Let us mention a few other restriction results in the number theoretic literature. Green and Tao [17] prove a restriction theorem for certain sets that are well controlled by the sieve, by a neat argument that is something of a hybrid between the approaches of Bourgain [3] and Green [16]. Recently, Keil [24, 25] has used another hybrid approach to study solutions of quadratic systems and forms in dense variables. He decomposes the counting function corresponding to his quadratic system, but does this based on a major arc decomposition on the Fourier side, rather than with sieves. This again appears to require sharp major arc estimates on a wide range. We also note the techniques of *pruning*, which are used to great effect in the circle method when limited major arc

information is available. In fact, results like Brüdern’s pruning lemma (see Lemma 2 of [8]) seem notably close to results from restriction theory, such as the technique of Bourgain [3] that we quote as Harmonic Analysis Result 2, below. But to apply pruning one generally needs at least some exponential sums around for which good major arc information is available, and in our problems we have none.

Finally turning to applications, by combining Theorems 1 and 2 with major arc estimates of Drappeau [12] we can readily handle ternary additive problems involving smooth numbers. As a canonical example, we obtain the following result.

Corollary 1. *There exists a large absolute constant $K > 0$ such that, for any large $\log^K x \leq y \leq x$, (and writing $u := (\log x)/\log y$),*

$$\#\{(a, b, c) \in \mathcal{S}(y)^3 : a, b, c \leq x, a + b = c\} = \frac{\Psi(x, y)^3}{2x} \left(1 + O\left(\frac{\log(u+1)}{\log y}\right) \right).$$

The error term $\log(u+1)/\log y$ tends to zero if $(\log y)/\log \log x \rightarrow \infty$. For smaller y (still satisfying $y \geq \log^K x$) one can obtain an error term tending to zero by replacing the main term $\Psi(x, y)^3/2x$ by something a little more complicated, but still explicit.

Assuming the Generalised Riemann Hypothesis, Lagarias and Soundararajan [28, 29] (see also [11]) have previously proved Corollary 1 for all $\log^{8+\epsilon} x \leq y \leq e^{\log^{1/2-\epsilon} x}$. There is analogous work of Ha [18] in the function field setting. Lagarias and Soundararajan investigated this equation as a close analogue of the *abc* conjecture, with the smoothness bound y taking the place of the radical of *abc* in that conjecture. Corollary 1 now gives the first unconditional proof that Lagarias and Soundararajan’s “*xyz*-smoothness exponent” is at most K , and in particular is finite².

Previously, the best known unconditional result like Corollary 1 was due to Drappeau [12], who could handle the range $e^{c\sqrt{\log x} \log \log x} \leq y \leq x$, improving on La Bretèche and Granville [6] who could handle $e^{\log^{2/3+\epsilon} x} \leq y \leq x$. We also mention an older result of Balog and A. Sárközy [2], who showed the existence of solutions to the equation $a + b + c = N$ with a, b, c all $e^{3\sqrt{\log N \log \log N}}$ -smooth, for any large N . They used exponential sums and the circle method, but with weights that allowed stronger estimates at the cost of losing an asymptotic for the number of solutions. G. Sárközy [32] improved this a bit, replacing the constant 3 in the exponent by $\sqrt{3/2} + \epsilon$, using the large sieve.

Our second application is to additive combinatorics, where we combine Theorems 1 and 2, major arc estimates, and the transference machinery of Green [16] to prove a version of Roth’s theorem inside the smooth numbers.

²Strictly speaking, Lagarias and Soundararajan need a, b, c to be coprime, but a simple inclusion-exclusion argument as in §8 of their paper [29] can be used to impose that condition.

Corollary 2. *For any $\beta > 0$ there exists a large constant $K = K(\beta) > 0$ such that the following is true. For any x that is large enough in terms of β , any $\log^K x \leq y \leq x$, and any set $B \subseteq \mathcal{S}(y) \cap [1, x]$ such that $\#B \geq \beta\Psi(x, y)$, there exist integers b and $d \neq 0$ such that*

$$b, b + d, b + 2d \in B.$$

As we mentioned, $\Psi(x, \log^K x) = x^{1-1/K+o(1)}$ as $x \rightarrow \infty$, so Corollary 2 exhibits a very sparse set inside which Roth's theorem holds. So far as the author knows, the only comparably sparse and non-trivial deterministic sets inside which Roth's theorem is known come from recent work of Mirek [30], who handled sets like the Piatetski–Shapiro primes (that is primes of the form $\lfloor n^{1/\gamma} \rfloor$, where γ is sufficiently close to 1 and $\lfloor \cdot \rfloor$ denotes integer part). In those cases the form of the function $n^{1/\gamma}$ again lets one obtain strong pointwise exponential sum bounds, unlike in our problem. More generally, Kohayakawa, Rödl, Schacht and Skokan [27] have shown graph-theoretically that if one has strong pointwise exponential sum bounds for a set then one can prove Roth's theorem inside that set, although they don't apply their condition to any particular cases. See the work of Conlon, Fox and Zhao [9] for the state of the art for progressions of length four or more. Despite this striking paucity of explicit examples, Kohayakawa, Łuczak and Rödl [26] proved that, with probability tending to 1 as $x \rightarrow \infty$, Roth's theorem holds inside random subsets of $[1, x]$ of cardinality $x^{1/2+o(1)}$ (roughly speaking).

To prove a result like Corollary 2, one needs a sharp restriction estimate for some $2 < p < 3$ for exponential sums over the set of interest, in this case the smooth numbers; and one needs to find a function $\nu(n)$ that upper bounds the characteristic function of that set, whose total sum is within a constant factor of the size of the set, *and* such that all exponential sums $\sum \nu(n)e(an/N)$ at certain well spaced points a/N are small. We give a precise statement of what is required as Transference Principle 1, in §6. In our case, Theorem 2 immediately provides the restriction estimate, and the obvious choice of $\nu(n)$ as the characteristic function $\mathbf{1}_{n \in \mathcal{S}(y)}$ turns out to be acceptable.

The reader might object that Corollary 2 doesn't really prove Roth's theorem inside the $\log^K x$ -smooth numbers for any fixed K , because K is required to grow with β . One could prove Corollary 2 with a large fixed K , but the proof would become more complicated because one would need to compensate for some irregularities of distribution of very smooth numbers, both p -adically (they are disproportionately unlikely to be coprime to small primes) and in an Archimedean sense (there are more $\log^K x$ -smooth numbers between 1 and $x/2$ than between $x/2$ and x , say). One could handle this using an analogue of Green's [16] *W-trick*, in which he restricted to working inside certain arithmetic progressions to overcome the irregular distribution of primes to small moduli (they are disproportionately likely to be coprime to the modulus). However, we have chosen not to present such an argument to avoid making this paper even longer.

We end this very long introduction by explaining the organisation of the rest of the paper. In §2 we state a few background results we shall need, concerning the distribution of smooth numbers, major arc estimates for smooth numbers, and a couple of harmonic analysis facts. Most of these are imported from the literature, but we need to prove a satisfactory upper bound for smooth numbers in segments of arithmetic progressions, and we also need to extend some known major arc estimates a little. Proofs of the new parts of major arc estimates are deferred to the appendix. In §3 we prove Theorem 1, and in §4 we prove Theorem 2. Finally, in §§5 – 6 we deduce our two corollaries.

2. SOME TOOLS

2.1. General results on smooth numbers. In this subsection we state some general smooth number estimates, but not exponential sum estimates, that we shall need.

We begin with a celebrated result of Hildebrand and Tenenbaum [22], that gives an asymptotic for $\Psi(x, y) := \sum_{n \leq x, n \in \mathcal{S}(y)} 1$ on a very wide range of x, y , in terms of a *saddle point* $\alpha = \alpha(x, y)$. We won't actually need this result much, but the subsequent result (which we will use extensively) will make much more sense if the reader is familiar with Hildebrand and Tenenbaum's asymptotic.

Smooth Numbers Result 1 (Hildebrand and Tenenbaum, 1986). *We have uniformly for $x \geq y \geq 2$,*

$$\Psi(x, y) = \frac{x^\alpha \zeta(\alpha, y)}{\alpha \sqrt{2\pi(1 + (\log x)/y) \log x \log y}} \left(1 + O\left(\frac{1}{\log(u+1)} + \frac{1}{\log y} \right) \right),$$

where $u = (\log x)/\log y$, $\zeta(s, y) := \sum_{n: n \text{ is } y \text{ smooth}} 1/n^s = \prod_{p \leq y} (1 - p^{-s})^{-1}$ for $\Re(s) > 0$, and $\alpha = \alpha(x, y) > 0$ is defined by

$$\sum_{p \leq y} \frac{\log p}{p^\alpha - 1} = \log x.$$

Hildebrand and Tenenbaum [22] also established a simple approximation for $\alpha(x, y)$ on the whole range $2 \leq y \leq x$. Their Lemma 2 implies, in particular, that when $\log x < y \leq x$ one has

$$\alpha(x, y) = 1 - \frac{\log(u \log(u+1))}{\log y} + O\left(\frac{1}{\log y} \right). \quad (2.1)$$

We also remark that, by definition, $\alpha(x, y)$ is a decreasing function of x for any fixed y .

To orient an unfamiliar reader, we also provide a more explicit estimate for $\Psi(x, y)$ on a slightly smaller range. Hildebrand [21] showed that

$$\Psi(x, y) = x \rho(u) \left(1 + O_\epsilon \left(\frac{\log(u+1)}{\log y} \right) \right), \quad e^{(\log \log x)^{5/3+\epsilon}} \leq y \leq x,$$

where the Dickman function $\rho(u)$ is a certain continuous function that satisfies $\rho(u) = e^{-(1+o(1))u \log u}$ as $u \rightarrow \infty$. Thus the y -smooth numbers are very sparse when $u = (\log x)/\log y$ is large³. In particular, when y decreases below $e^{\sqrt{(1/2) \log x \log \log x}}$ there is a change in that $\Psi(x, y)/x \approx \rho(u)$ becomes smaller than $1/y$. As major arc estimates for smooth numbers are valid when $q \leq y^c$, this explains why a classical approach to ternary additive problems with smooth numbers fails for y smaller than about $e^{\sqrt{(1/2) \log x \log \log x}}$.

The next result will be very important to our arguments. It is a neat “local” result of La Bretèche and Tenenbaum [7], that compares $\Psi(x, y)$ with $\Psi(x/d, y)$ on a very wide range of the parameters.

Smooth Numbers Result 2 (See Théorème 2.4(i) of La Bretèche and Tenenbaum [7]). *Let $2 \leq y \leq x$, and suppose that $d \geq 1$. Then we have*

$$\Psi(x/d, y) \ll \frac{1}{d^\alpha} \Psi(x, y),$$

where $\alpha = \alpha(x, y)$ denotes the saddle-point corresponding to the y -smooth numbers less than x .

As mentioned in the introduction, we will also need an upper bound for the quantity of smooth numbers in a segment of an arithmetic progression. By combining Smooth Numbers Result 2 with a simple but powerful method of Friedlander [14], we will prove the following result.

Smooth Numbers Result 3. *Let $\log X \leq y \leq X$ be large, and suppose $q \geq 1$ and $qy \leq Z \leq X$. Then*

$$\sum_{\substack{X \leq n \leq X+Z, \\ n \equiv a \pmod q, \\ n \in \mathcal{S}(y)}} 1 \ll \frac{Z}{qX} \Psi(X, y) \left(\frac{Xq}{Z} \right)^{1-\alpha} \log X,$$

where $\alpha = \alpha(X, y)$ denotes the saddle-point corresponding to the y -smooth numbers less than X .

Note that if the y -smooth numbers were roughly equidistributed in short intervals and arithmetic progressions one might expect an upper bound $\ll (Z/qX) \Psi(X, y)$, and if y is at least a large power of $\log X$ then $1 - \alpha$ will be “close” to 0, so the bound in Smooth Numbers Result 3 will be of roughly the correct shape (although a more precise bound would certainly be desirable). One could reduce or remove the factor $\log X$, and weaken the assumptions on y and Z , with more work and by slightly reformulating the

³When $\log^{1+\epsilon} x \leq y \leq x$, Theorem 2(ii) and Corollary 2 of Hildebrand and Tenenbaum [22] also imply that $\Psi(x, y) = x\rho(u)e^{O_\epsilon(u)}$. This implies the estimate $\Psi(x, \log^K x) = x^{1-1/K+o(1)}$ for any fixed $K > 1$, as $x \rightarrow \infty$, which we mentioned several times in the introduction (and which can actually be proved more simply, and is true when $K = 1$ as well).

conclusion, but this will be unnecessary for us. Let us also remark that Friedlander [14] proved an upper bound roughly comparable to Smooth Numbers Result 3 for $y \geq e^{\log^{4/5} X}$ (and for $Z = X$), and later Balog and Pomerance [1] used his method to handle $y \geq \exp\{(\log \log X)^2\}$, although their bound was weaker (primarily since they did not have access to a suitable result like Smooth Numbers Result 2).

Turning to the proof of Smooth Numbers Result 3, suppose at first (for simplicity) that $(a, q) = 1$. Write $P(m)$ for the largest prime factor of the integer m , and write $p(n)$ for the smallest prime factor of the integer n . Then we see

$$\sum_{\substack{X \leq n \leq X+Z, \\ n \equiv a \pmod q, \\ n \in \mathcal{S}(y)}} 1 = \sum_{\substack{Z/xy < m \leq Z/q, \\ m/P(m) \leq Z/xy, \\ m \in \mathcal{S}(y), \\ (m, q) = 1}} \sum_{\substack{X/m \leq n \leq (X+Z)/m, \\ p(n) \geq P(m), \\ mn \equiv a \pmod q, \\ n \in \mathcal{S}(y)}} 1,$$

since every smooth number from the interval $[X, X + Z]$ can be written uniquely as such a product mn , by revealing its prime factors one at a time, starting with the smallest, until the product exceeds Z/xy . Note that we may impose the condition that $(m, q) = 1$ since otherwise the condition $mn \equiv a \pmod q$ cannot possibly be satisfied, given our assumption that $(a, q) = 1$.

Next, if we write m^{-1} to mean the multiplicative inverse of m modulo q , and relax the inner summation, we obtain

$$\sum_{\substack{X \leq n \leq X+Z, \\ n \equiv a \pmod q, \\ n \in \mathcal{S}(y)}} 1 \leq \sum_{\substack{Z/xy < m \leq Z/q, \\ m/P(m) \leq Z/xy, \\ m \in \mathcal{S}(y), \\ (m, q) = 1}} \sum_{\substack{X/m \leq n \leq (X+Z)/m, \\ n \equiv am^{-1} \pmod q}} 1 \leq \frac{2Z}{q} \sum_{\substack{Z/xy < m \leq Z/q, \\ m/P(m) \leq Z/xy, \\ m \in \mathcal{S}(y), \\ (m, q) = 1}} \frac{1}{m},$$

in view of the fact that $Z/m \geq q$. Dropping the restrictions that $m/P(m) \leq Z/xy$ and $(m, q) = 1$, we see

$$\begin{aligned} \sum_{\substack{Z/xy < m \leq Z/q, \\ m \in \mathcal{S}(y)}} \frac{1}{m} &\leq \sum_{0 \leq j \leq (\log y)/\log 2} \frac{1}{2^j(Z/xy)} \sum_{\substack{2^j(Z/xy) \leq m \leq 2^{j+1}(Z/xy), \\ m \in \mathcal{S}(y)}} 1 \\ &\leq \sum_{0 \leq j \leq (\log y)/\log 2} \frac{1}{2^j(Z/xy)} \Psi\left(\frac{2^{j+1}Z}{Xxy} X, y\right) \\ &\ll \Psi(X, y) \left(\frac{Z}{Xxy}\right)^\alpha \sum_{0 \leq j \leq (\log y)/\log 2} \frac{2^{j(\alpha-1)}}{(Z/xy)}, \end{aligned}$$

in view of Smooth Numbers Result 2. Finally, if we write $u = (\log X)/\log y$ then, since $\alpha = 1 - \frac{\log(u \log(u+1)) + O(1)}{\log y}$ when $\log X \leq y \leq X$, and since $y^{1-\alpha} = O(u \log(u+1))$, we

certainly have

$$\sum_{\substack{Z/qy < m \leq Z/q, \\ m \in \mathcal{S}(y)}} \frac{1}{m} \ll \frac{1}{X} \Psi(X, y) \left(\frac{Xqy}{Z} \right)^{1-\alpha} \frac{\log y}{\log(u+1)} \ll \frac{1}{X} \Psi(X, y) \left(\frac{Xq}{Z} \right)^{1-\alpha} \log X,$$

and Smooth Numbers Result 3 follows on putting everything together.

It only remains to remove our simplifying assumption that $(a, q) = 1$. If $(a, q) = h$ for some h that is not y -smooth then the sum in Smooth Numbers Result 3 has no terms, so the result is trivial. If $(a, q) = h$ for some $h \neq 1$ that is y -smooth then

$$\sum_{\substack{X \leq n \leq X+Z, \\ n \equiv a \pmod{q}, \\ n \in \mathcal{S}(y)}} 1 = \sum_{\substack{X/h \leq n \leq (X+Z)/h, \\ n \equiv a/h \pmod{q/h}, \\ n \in \mathcal{S}(y)}} 1,$$

where $(a/h, q/h) = 1$, and where all the hypotheses of Smooth Numbers Result 3 are satisfied by the sum on the right. (Note that $y \leq X/q \leq X/h$, by assumption.) The result then follows by the above argument, and by using Smooth Numbers Result 2 again to obtain that $\Psi(X/h, y) \ll h^{-\alpha} \Psi(X, y)$.

Q.E.D.

2.2. Major arc estimates. In this subsection we state some major arc estimates for exponential sums over smooth numbers. We will need these both as an ingredient for the proof of Theorem 2, to avoid logarithmic losses, and when deducing Corollaries 1 and 2.

Major Arc Estimate 1. *There exist two absolute positive constants c_1, c_2 such that the following is true. For any large $\log^{c_1} x \leq y \leq x$, any $q \leq y^{c_2}$, any $(a, q) = 1$, any $\delta \in \mathbb{R}$, and any $A > 0$, we have*

$$\sum_{\substack{n \leq x, \\ n \in \mathcal{S}(y)}} e(n(a/q + \delta)) = V(x, y; q, \delta) + O_A(\Psi(x, y)(1 + |\delta x|)(y^{-c_2} + \frac{e^{-c_2 u / \log^2(u+1)}}{\log^A x})),$$

where $u := (\log x) / \log y$, $V(x, y; q, \delta) := \sum_{\substack{n \leq x, \\ n \in \mathcal{S}(y)}} \frac{\mu(q/(q, n))}{\phi(q/(q, n))} e(n\delta)$, and μ, ϕ denote the Möbius and Euler totient functions.

Major Arc Estimate 1 reduces the study of exponential sums over smooth numbers on the major arcs to an investigation of the main term $V(x, y; q, \delta)$. The estimate more or less follows from Proposition 2.2 of Drappeau [12], with the choice $Q := \min\{y^{c_2}, e^{c_2 \sqrt{\log x}}\}$, except that proposition includes an extra term $W(x, y; q, \delta)$ reflecting the possible existence of an exceptional (Siegel) zero. Given our current knowledge about character sums over smooth numbers, it actually isn't too difficult to bound $W(x, y; q, \delta)$ satisfactorily on the complete range in Major Arc Estimate 1. We

supply the relevant argument, which combines with Proposition 2.2 of Drappeau [12] to prove Major Arc Estimate 1, in the appendix.

We will of course also need some information about $V(x, y; q, \delta)$.

Major Arc Estimate 2. *There exist two absolute positive constants c_1, c_2 such that the following is true. For any large $\log^{c_1} x \leq y \leq x$, any $q \leq y^{1/4}$, and any $|\delta| \leq \min\{y^{c_2}/x, e^{c_2 \log^{1/4} x}/x\}$, we have*

$$V(x, y; q, \delta) = \frac{\Psi(x, y)}{x} V(x, x; q, \delta) + O\left(\Psi(x, y) \frac{\log(u+1)}{\log y} \frac{2^{\omega(q)} q^{1-\alpha} \log^2(q+1) \log^3(2+|\delta x|)}{\phi(q) (1+|\delta x|)^\alpha}\right),$$

where $\omega(q)$ denotes the number of distinct prime factors of q , and $\alpha = \alpha(x, y)$ denotes the saddle-point corresponding to the y -smooth numbers less than x .

When $\log^{c_1} x \leq y \leq e^{\sqrt{\log x}}$, say, Major Arc Estimate 2 follows from Proposition 2.3 of Drappeau [12] (in fact with a more precise error term, and on a wider range of q and δ), except that the main term in Drappeau's result is expressed in terms of a certain two variable Mellin transform. It is more convenient for most applications to have a main term $(\Psi(x, y)/x)V(x, x; q, \delta)$, which will allow us simply to compare the exponential sum over smooth numbers with the complete sum, and in deducing that the error term is necessarily degraded a bit. When y is larger, Drappeau's error term is less precise than in Major Arc Estimate 2 for a significant but technical reason, but we can instead deduce the result from an older estimate of La Bretèche [5]. One could presumably also use Proposition 2.4 of Drappeau [12], but the computations needed seem more formidable. We give details of all these deductions in the appendix.

2.3. Other analytic tools. In this subsection we state two general harmonic analysis results that we shall need for the proof of Theorem 2.

The first result is a very well known inequality that gives information about the (equi)distribution of a sequence in terms of bounds on its exponential sums.

Harmonic Analysis Result 1 (Erdős–Turán inequality, 1948, see Corollary 1.1 of Montgomery [31]). *Let $0 \leq u_1, u_2, \dots, u_N \leq 1$ be any points. Then for any $J \in \mathbb{N}$, and any $\alpha \leq \beta \leq \alpha + 1$, we have*

$$|\#\{1 \leq n \leq N : u_n \in [\alpha, \beta] \bmod 1\} - (\beta - \alpha)N| \leq \frac{N}{J+1} + 3 \sum_{j=1}^J \frac{1}{j} \left| \sum_{n=1}^N e(ju_n) \right|.$$

The second result is a tool from Bourgain's [3] restriction argument, which essentially encodes the fact that major arcs are fairly "simple" or "well spaced". See Lemma 2 of Brüdern [8] for a closely related estimate.

Harmonic Analysis Result 2 (See pp 305-307 of Bourgain [3]). *Let x be large, let $Q \geq 1$ and $1/x \leq v \leq 1/2$ be any parameters, and let $\epsilon > 0$ and $A > 0$ be arbitrary. Define*

$$G(\theta) := \sum_{q \leq Q} \frac{1}{q} \sum_{a=0}^{q-1} \frac{\mathbf{1}_{\|\theta - a/q\| \leq v}}{1 + x\|\theta - a/q\|},$$

where as usual $\|\cdot\|$ denotes distance to the nearest integer, and $\mathbf{1}$ denotes the indicator function.

Then if $\theta_1, \dots, \theta_R$ are any real numbers such that $\|\theta_r - \theta_s\| \geq 1/x$ when $r \neq s$,

$$\sum_{1 \leq r, s \leq R} G(\theta_r - \theta_s) \ll_{\epsilon, A} RQ^\epsilon \log(1 + vx) + \frac{R^2 Q \log(1 + vx)}{x} + \frac{R^2 \log(1 + vx)}{Q^A}.$$

To orient the reader, we note that we shall apply Harmonic Analysis Result 2 towards the end of the proof of Theorem 2 (once a function like $G(\theta)$ has appeared), and in a situation where Q is large compared with some other parameters in the argument, but fairly small compared with x . Thus the term $RQ^\epsilon \log(1 + vx)$, which basically reflects the diagonal contribution from terms $r = s$, will be the dominant one, and this nice behaviour will finally prove Theorem 2.

3. PROOF OF THEOREM 1

For simplicity of writing, in this section we set $L := 2(1 + |\delta x|)$. Note that the conditions of Theorem 1 imply that $q^2 y^3 L^2 \leq x$. Recall also that we write $P(m)$ for the largest prime factor of the integer m , and write $p(n)$ for the smallest prime factor of the integer n .

To prove Theorem 1, we begin with a standard move of decomposing $\sum_{\substack{n \leq x, \\ n \in \mathcal{S}(y)}} e(n\theta)$ into multiple sums, similarly as in the proof of Smooth Numbers Result 3. We see

$$\begin{aligned} \sum_{\substack{n \leq x, \\ n \in \mathcal{S}(y)}} e(n\theta) &= \sum_{\substack{x/qL \leq n \leq x, \\ n \in \mathcal{S}(y)}} e(n\theta) + O(\Psi(x/qL, y)) \\ &= \sum_{\substack{qLy < m \leq qLy^2, \\ m/P(m) \leq qLy, \\ m \in \mathcal{S}(y)}} \sum_{\substack{x/qLm \leq n \leq x/m, \\ p(n) \geq P(m), \\ n \in \mathcal{S}(y)}} e(mn\theta) + O(\Psi(x/qL, y)), \end{aligned}$$

noting that $qLy^2 < x/qL$, by hypothesis, so certainly every y -smooth number from the interval $[x/qL, x]$ has a unique decomposition in the form mn (by taking m to consist of the smallest prime factors of the number).

We wish to continue by applying the Cauchy–Schwarz inequality, which will allow us to complete the sum over m so that we can perform estimations. However, in order to do this we need to remove the dependence between n and m in the condition $p(n) \geq P(m)$, and to do it efficiently we would also like the m variable to run over dyadic ranges.

Thus we write

$$\begin{aligned} \sum_{\substack{n \leq x, \\ n \in \mathcal{S}(y)}} e(n\theta) &= \sum_{0 \leq j \leq \frac{\log y}{\log 2}} \sum_{p \leq y} \sum_{\substack{2^j qLy < m \leq qLy \cdot \min\{2^{j+1}, p\}, \\ P(m)=p}} \sum_{\substack{x/qLm \leq n \leq x/m, \\ p(n) \geq p, \\ n \in \mathcal{S}(y)}} e(mn\theta) \\ &\quad + O(\Psi(x/qL, y)), \end{aligned}$$

noting that if $P(m) = p$ then the condition that m be y -smooth is automatically satisfied, and the condition that $m/P(m) \leq qLy$ can be written as $m \leq qLyp$.

Now let $0 \leq j \leq (\log y)/\log 2$ be fixed. For simplicity of writing, we will let \sum_m denote $\sum_{\substack{2^j qLy < m \leq qLy \cdot \min\{2^{j+1}, p\}, \\ P(m)=p}}$. Let us also observe that we may restrict the sum over primes to $p \geq 2^j$, since otherwise the sum over m is empty. (This small refinement is not very important, but will lead to a slightly neater and stronger bound at the end.) Then using the Cauchy–Schwarz inequality, we see

$$\begin{aligned} \left| \sum_{p \leq y} \sum_m \sum_{\substack{x/qLm \leq n \leq x/m, \\ p(n) \geq p, \\ n \in \mathcal{S}(y)}} e(mn\theta) \right| &\leq \sqrt{\left(\sum_{p \leq y} \sum_m 1 \right) \left(\sum_{p \leq y} \sum_m \left| \sum_{\substack{x/qLm \leq n \leq x/m, \\ p(n) \geq p, \\ n \in \mathcal{S}(y)}} e(mn\theta) \right|^2 \right)} \\ &\leq \sqrt{\Psi(2^{j+1}qLy, y)} \sqrt{\sum_{2^j \leq p \leq y} \sum_{\substack{2^j qLy < m' \leq \frac{qLy}{p} \min\{2^{j+1}, p\} \\ x/(qLm'p) \leq n \leq x/(m'p), \\ p(n) \geq p, \\ n \in \mathcal{S}(y)}} \left| \sum_{\substack{x/qLm'p \leq n \leq x/(m'p), \\ p(n) \geq p, \\ n \in \mathcal{S}(y)}} e(m'pn\theta) \right|^2} \\ &\ll \sqrt{\Psi(2^{j+1}qLy, y)} \sqrt{\sum_{2^j \leq p \leq y} \sum_{\substack{n_1, n_2 \leq \frac{x}{2^j qLy}, \\ p(n_1), p(n_2) \geq p, \\ n_1, n_2 \in \mathcal{S}(y)}} \min \left\{ \frac{2^{j+1}qLy}{p}, \frac{1}{\|(n_1 - n_2)p\theta\|} \right\}}, \end{aligned}$$

where as usual $\|\cdot\|$ denotes distance to the nearest integer, and the final inequality follows by expanding the square and summing the geometric progression over m' .

Now there is an unfortunate but standard complication, in that we must distinguish cases according as a prime p does or does not divide q (where $\theta = a/q + \delta$). Recall that $L := 2(1 + |\delta x|)$. If $p \nmid q$, and if $n_1, n_2 \leq x/2^j qLy$, then we see

$$\|(n_1 - n_2)p\theta\| = \left\| \frac{(n_1 - n_2)pa}{q} + (n_1 - n_2)p\delta \right\| \asymp \begin{cases} \left\| \frac{bpa}{q} \right\| & \text{if } n_1 - n_2 \equiv b \pmod{q}, \text{ for } b \neq 0 \\ |(n_1 - n_2)p\delta| & \text{if } n_1 - n_2 \equiv 0 \pmod{q}, \end{cases}$$

since by definition of L we have $|(n_1 - n_2)p\delta| \leq p|\delta x|/2^j qLy \leq |\delta x|/qL \leq 1/2q$. This inequality is one of the reasons for setting the ranges of n and m as we did. If instead

$p \mid q$, then we similarly have

$$|(n_1 - n_2)p\theta| = \left| \frac{(n_1 - n_2)a}{q/p} + (n_1 - n_2)p\delta \right| \asymp \begin{cases} \left| \frac{ba}{q/p} \right| & \text{if } n_1 - n_2 \equiv b \pmod{\frac{q}{p}}, \text{ for } b \neq 0 \\ |(n_1 - n_2)p\delta| & \text{if } n_1 - n_2 \equiv 0 \pmod{\frac{q}{p}}. \end{cases}$$

So to summarise our work so far, if we define

$$\mathcal{T}_j(r) := \max_{1 \leq b \leq r-1} \sum_{\substack{n_1, n_2 \leq \frac{x}{2^j q L y}, \\ n_1, n_2 \in \mathcal{S}(y), \\ n_1 - n_2 \equiv b \pmod{r}}} 1,$$

(suppressing mention of the quantities x, y, q, L on which $\mathcal{T}_j(r)$ of course also depends, and throwing away the condition that $p(n_1), p(n_2) \geq p$), then we deduce that

$$\left| \sum_{\substack{n \leq x, \\ n \in \mathcal{S}(y)}} e(n\theta) \right| \ll \Psi(x/qL, y) + \sum_{0 \leq j \leq \frac{\log y}{\log 2}} \sqrt{\Psi(2^{j+1}qLy, y)} \sqrt{S_1} + \sum_{0 \leq j \leq \frac{\log y}{\log 2}} \sqrt{\Psi(2^{j+1}qLy, y)} \sqrt{S_2}, \quad (3.1)$$

where

$$S_1 = S_1(j) := \sum_{\substack{2^j \leq p \leq y, \\ p \nmid q}} \mathcal{T}_j(q) \sum_{b=1}^{q-1} \min \left\{ \frac{2^{j+1}qLy}{p}, \frac{q}{b} \right\} + \sum_{\substack{2^j \leq p \leq y, \\ p \mid q}} \mathcal{T}_j(q/p) \sum_{b=1}^{(q/p)-1} \min \left\{ \frac{2^{j+1}qLy}{p}, \frac{q}{pb} \right\},$$

and where

$$\begin{aligned} S_2 = S_2(j) &:= \sum_{\substack{2^j \leq p \leq y, \\ p \nmid q}} \frac{1}{p} \sum_{\substack{n_1, n_2 \leq \frac{x}{2^j q L y}, \\ n_1, n_2 \in \mathcal{S}(y), \\ n_1 - n_2 \equiv 0 \pmod{q}}} \min \left\{ 2^{j+1}qLy, \frac{1}{|(n_1 - n_2)\delta|} \right\} \\ &+ \sum_{\substack{2^j \leq p \leq y, \\ p \mid q}} \frac{1}{p} \sum_{\substack{n_1, n_2 \leq \frac{x}{2^j q L y}, \\ n_1, n_2 \in \mathcal{S}(y), \\ n_1 - n_2 \equiv 0 \pmod{q/p}}} \min \left\{ 2^{j+1}qLy, \frac{1}{|(n_1 - n_2)\delta|} \right\}. \end{aligned}$$

Here we used the fact that, if $(a, q) = 1$ and $p \nmid q$, the numbers bpa vary over all non-zero residue classes modulo q as the numbers b vary over all non-zero residue classes (though not in general in the same order, of course), similarly for the numbers ba modulo q/p in the case where $p \mid q$.

Thus far we haven't deployed much information about the y -smooth numbers, apart from their good factorisation properties. To complete the proof of Theorem 1 we need to bound the sums $\mathcal{T}_j(q), \mathcal{T}_j(q/p)$, and the other remaining sums over n_1, n_2 inside S_2 , and in doing this we shall use Smooth Numbers Results 2 and 3. In fact we can show:

Proposition 1. *Let $\log^{1.1} x \leq y \leq x^{1/3}$, let $q \geq 1$ and $L = 2(1 + |\delta x|)$ be as above, and suppose that $q^2 y^3 L^2 \leq x$. Then for any $0 \leq j \leq (\log y)/\log 2$, and any prime $p \mid q$, we have*

$$\mathcal{T}_j(q) \ll \frac{\Psi(x/2^j qLy, y)^2}{q} q^{1-\alpha(x,y)} \log x \quad \text{and} \quad \mathcal{T}_j(q/p) \ll \frac{\Psi(x/2^j qLy, y)^2}{q/p} (q/p)^{1-\alpha(x,y)} \log x.$$

Under the same hypotheses, we have

$$\sum_{\substack{n_1, n_2 \leq \frac{x}{2^j qLy}, \\ n_1, n_2 \in \mathcal{S}(y), \\ n_1 - n_2 \equiv 0 \pmod{q}}} \min \left\{ 2^{j+1} qLy, \frac{1}{|(n_1 - n_2)\delta|} \right\} \ll 2^j y \Psi\left(\frac{x}{2^j qLy}, y\right)^2 (qL)^{1-\alpha(x,y)} \log x \log L,$$

and similarly for any prime $p \mid q$ we have

$$\sum_{\substack{n_1, n_2 \leq \frac{x}{2^j qLy}, \\ n_1, n_2 \in \mathcal{S}(y), \\ n_1 - n_2 \equiv 0 \pmod{q/p}}} \min \left\{ 2^{j+1} qLy, \frac{1}{|(n_1 - n_2)\delta|} \right\} \ll p 2^j y \Psi\left(\frac{x}{2^j qLy}, y\right)^2 \left(\frac{q}{p} L\right)^{1-\alpha(x,y)} \log x \log L.$$

Assuming the truth of Proposition 1 for a moment, we can quickly finish the proof of Theorem 1. Note that for any $p \leq y$ we have

$$\sum_{b=1}^{q-1} \min \left\{ \frac{2^{j+1} qLy}{p}, \frac{q}{b} \right\} = \sum_{b=1}^{q-1} \frac{q}{b} \ll q \log q, \quad \text{and} \quad \sum_{b=1}^{(q/p)-1} \min \left\{ \frac{2^{j+1} qLy}{p}, \frac{q}{pb} \right\} \ll (q/p) \log q,$$

so we deduce from (3.1) and Proposition 1 that $\left| \sum_{\substack{n \leq x, \\ n \in \mathcal{S}(y)}} e(n\theta) \right|$ is

$$\begin{aligned} &\ll \Psi(x/qL, y) + \sum_{0 \leq j \leq \frac{\log y}{\log 2}} \sqrt{\Psi(2^{j+1} qLy, y)} \sqrt{\frac{y}{\log y} \log q \Psi\left(\frac{x}{2^j qLy}, y\right)^2 q^{1-\alpha} \log x} \\ &\quad + \sum_{0 \leq j \leq \frac{\log y}{\log 2}} \sqrt{\Psi(2^{j+1} qLy, y)} \sqrt{\left(\sum_{\substack{2^j \leq p \leq y, \\ p \nmid q}} \frac{1}{p} + \sum_{\substack{2^j \leq p \leq y, \\ p \mid q}} 1 \right) 2^j y \Psi\left(\frac{x}{2^j qLy}, y\right)^2 (qL)^{1-\alpha} \log x \log L}. \end{aligned}$$

Next we use the crude upper bound $\Psi(2^{j+1} qLy, y) \leq 2^{j+1} qLy$, (which, however, shouldn't lose much unless qL is rather large, in which case we already gain a lot elsewhere), together with the bounds $\Psi(x/qL, y) \ll \Psi(x, y)/(qL)^\alpha$ and $\Psi(x/2^j qLy, y) \ll \Psi(x, y)/(2^j qLy)^\alpha$ coming from Smooth Numbers Result 2, and the elementary estimate

$$\sum_{\substack{2^j \leq p \leq y, \\ p \nmid q}} \frac{1}{p} + \sum_{\substack{2^j \leq p \leq y, \\ p \mid q}} 1 \ll \sum_{p \leq y} \frac{\log p}{(j+1)p} + \sum_{\substack{p \mid q, \\ p \geq 2^j}} 1 \ll \frac{\log(qy)}{j+1}.$$

Inserting these bounds, we obtain that

$$\left| \sum_{\substack{n \leq x, \\ n \in \mathcal{S}(y)}} e(n\theta) \right| \ll \frac{\Psi(x, y)}{(qL)^\alpha} + \sum_{0 \leq j \leq \frac{\log y}{\log 2}} \frac{\Psi(x, y)}{\sqrt{2^j qL}} \sqrt{\frac{1}{\log y} \log q (2^j qLy)^{2(1-\alpha)} q^{1-\alpha} \log x} \\ + \frac{\Psi(x, y)}{\sqrt{qL}} \sum_{0 \leq j \leq \frac{\log y}{\log 2}} \sqrt{\frac{\log(qy)}{j+1} (2^j qLy)^{2(1-\alpha)} (qL)^{1-\alpha} \log x \log L}.$$

In Theorem 1 we have $y \geq \log^3 x$, and so $\alpha(x, y) \geq 2/3 + o(1)$ (by the approximation (2.1)) and so the first term is more than good enough for the theorem. The second term is visibly smaller than the third, and the third term is

$$\ll \frac{\Psi(x, y)}{\sqrt{qL}} (qL)^{(3/2)(1-\alpha)} \sqrt{\log(qy) \log x \log L} \sum_{0 \leq j \leq \frac{\log y}{\log 2}} \frac{2^{j(1-\alpha)} y^{1-\alpha}}{\sqrt{j+1}} \\ \ll \frac{\Psi(x, y)}{\sqrt{qL}} (qL)^{(3/2)(1-\alpha)} \sqrt{\log(qy) \log x \log L} \left(y^{(3/2)(1-\alpha)} \sqrt{\log y} + \sum_{\frac{\log y}{2 \log 2} \leq j \leq \frac{\log y}{\log 2}} \frac{2^{j(1-\alpha)} y^{1-\alpha}}{\sqrt{\log y}} \right) \\ \ll \frac{\Psi(x, y)}{\sqrt{qL}} (qL)^{(3/2)(1-\alpha)} \sqrt{\log(qy) \log x \log L} \sqrt{\log y} \frac{y^{2(1-\alpha)}}{1 + |1 - \alpha| \log y}.$$

Here the final inequality uses the bound $\sum_{j \leq \frac{\log y}{\log 2}} 2^{j(1-\alpha)} \ll y^{1-\alpha} \min\{\log y, 1/|1 - \alpha|\}$, which follows from summing the geometric progression. Then since we have $\alpha = 1 - \frac{\log(u \log u) + O(1)}{\log y}$ and $y^{1-\alpha} = O(u \log u)$, by (2.1), the conclusion of Theorem 1 immediately follows on remembering that $\sqrt{\log x \log y} = (\log x)/\sqrt{u}$.

Q.E.D.

Now it only remains to prove Proposition 1. However, by definition we have

$$\mathcal{T}_j(q) = \max_{1 \leq b \leq q-1} \sum_{\substack{n_1 \leq x/2^j qLy, \\ n_1 \in \mathcal{S}(y)}} \sum_{\substack{n_2 \leq x/2^j qLy, \\ n_2 \equiv n_1 - b \pmod{q}, \\ n_2 \in \mathcal{S}(y)}} 1,$$

and by applying Smooth Numbers Result 3 on dyadic intervals $[X, 2X]$, beginning with $X = qy$ and continuing with X of the form $2^i qy$, we see that the inner sum is

$$\ll \sum_{\substack{n_2 \leq qy, \\ n_2 \equiv n_1 - b \pmod{q}, \\ n_2 \in \mathcal{S}(y)}} 1 + \sum_{X=2^i qy \leq x/2^j qLy} \frac{\Psi(X, y)}{q} q^{1-\alpha(X, y)} \log X \ll y + \frac{\Psi(x/2^j qLy, y)}{q} q^{1-\alpha(x, y)} \log x.$$

Here we bounded the first sum trivially, and we used Smooth Numbers Result 2 to obtain that $\sum_{X=2^i qy \leq x/2^j qLy} \Psi(X, y) \ll \Psi(x/2^j qLy, y)$. Let us also note that

$$\Psi(x/2^j qLy, y) \gg \Psi(x/2^j q^2 Ly, y) q^\alpha \geq \Psi(Ly, y) q^\alpha \geq y q^\alpha,$$

in view of Smooth Numbers Result 2 and then our assumption that $x \geq q^2 y^3 L^2$. We deduce that the trivial first term $y \ll \frac{\Psi(x/2^j qLy, y)}{q} q^{1-\alpha} \log x$ as well, so we indeed have

$$\mathcal{T}_j(q) \ll \sum_{\substack{n_1 \leq x/2^j qLy, \\ n_1 \in \mathcal{S}(y)}} \frac{\Psi(x/2^j qLy, y)}{q} q^{1-\alpha} \log x = \frac{\Psi(x/2^j qLy, y)^2}{q} q^{1-\alpha} \log x,$$

as claimed in Proposition 1. The claimed bound $\mathcal{T}_j(q/p) \ll \frac{\Psi(x/2^j qLy, y)^2}{q/p} (q/p)^{1-\alpha} \log x$ follows in exactly the same way.

To prove the second pair of bounds in Proposition 1, for the double sums of the form $\sum_{\substack{n_1, n_2 \leq \frac{x}{2^j qLy}, \\ n_1, n_2 \in \mathcal{S}(y), \\ n_1 - n_2 \equiv 0 \pmod{q}}} \min \left\{ 2^{j+1} qLy, \frac{1}{|(n_1 - n_2)\delta|} \right\}$, we distinguish two cases. If $|\delta| \leq 1/x$ then we

have $L = 2(1 + |\delta x|) \asymp 1$, and the bounds can be proved exactly as above on upper bounding $\min \left\{ 2^{j+1} qLy, \frac{1}{|(n_1 - n_2)\delta|} \right\}$ by $2^{j+1} qLy \ll 2^j qy$.

The other case is where $|\delta| > 1/x$, and so we have $L \asymp |\delta x|$. For simplicity of writing, let us temporarily use \sum^\dagger to denote a sum over pairs of integers $n_1, n_2 \leq \frac{x}{2^j qLy}$ that are y -smooth, and satisfy $n_1 - n_2 \equiv 0 \pmod{q}$. Then in the first place we have

$$\sum_{|n_1 - n_2| \leq x/2^j qL^2 y}^\dagger \min \left\{ 2^{j+1} qLy, \frac{1}{|(n_1 - n_2)\delta|} \right\} \ll 2^j qLy \sum_{\substack{n_1 \leq \frac{x}{2^j qLy}, \\ n_1 \in \mathcal{S}(y)}} \sum_{\substack{n_2 - n_1 \leq x/2^j qL^2 y, \\ n_2 \in \mathcal{S}(y), \\ n_2 \equiv n_1 \pmod{q}}} 1,$$

and as above we have $x/2^j qL^2 y \geq qy$ because of our hypothesis that $x \geq q^2 y^3 L^2$, so Smooth Numbers Results 2 and 3 are applicable and imply this is all

$$\ll 2^j qLy \sum_{\substack{n_1 \leq \frac{x}{2^j qLy}, \\ n_1 \in \mathcal{S}(y)}} \frac{\Psi(x/2^j qLy, y)}{qL} (qL)^{1-\alpha} \log x = 2^j y \Psi(x/2^j qLy, y)^2 (qL)^{1-\alpha} \log x.$$

Similarly, for any $0 \leq r \leq (\log L)/\log 2$ we have

$$\begin{aligned} & \sum_{\substack{\frac{2^r x}{2^j qL^2 y} < |n_1 - n_2| \leq \frac{2^{r+1} x}{2^j qL^2 y}}}^\dagger \min \left\{ 2^{j+1} qLy, \frac{1}{|(n_1 - n_2)\delta|} \right\} \ll \frac{2^j qLy}{2^r} \sum_{\substack{n_1 \leq \frac{x}{2^j qLy}, \\ n_1 \in \mathcal{S}(y)}} \sum_{\substack{n_2 - n_1 \leq \frac{2^{r+1} x}{2^j qL^2 y}, \\ n_2 \in \mathcal{S}(y), \\ n_2 \equiv n_1 \pmod{q}}} 1 \\ & \ll \frac{2^j qLy}{2^r} \sum_{\substack{n_1 \leq \frac{x}{2^j qLy}, \\ n_1 \in \mathcal{S}(y)}} \Psi(x/2^j qLy, y) \left(\frac{2^r}{qL} \right)^\alpha \log x \\ & \ll 2^j y \Psi(x/2^j qLy, y)^2 (qL)^{1-\alpha} \log x, \end{aligned}$$

and the Proposition 1 bound for $\sum^\dagger \min \left\{ 2^{j+1} qLy, \frac{1}{|(n_1 - n_2)\delta|} \right\}$ follows by summing over r .

The claimed bound for $\sum_{\substack{n_1, n_2 \leq \frac{x}{2^j q Ly}, \\ n_1, n_2 \in \mathcal{S}(y), \\ n_1 - n_2 \equiv 0 \pmod{q/p}}} \min \left\{ 2^{j+1} q Ly, \frac{1}{|(n_1 - n_2)\delta|} \right\}$, losing a factor p^α because of the weaker congruence condition, follows in exactly the same way.

Q.E.D.

4. PROOF OF THEOREM 2

Throughout this section we shall assume that $y \leq x^{1/100}$, since if $x^{1/100} < y \leq x$ then $\Psi(x, y) \asymp x$, in which case Theorem 2 is an easy consequence of Parseval's identity.

To prove Theorem 2, we shall actually prove the following large values (or “distributional”) estimate for the number of $1/x$ -neighbourhoods on which $\left| \sum_{\substack{n \leq x, \\ n \in \mathcal{S}(y)}} a_n e(n\theta) \right|$ is large. Recall that $\|\cdot\|$ denotes distance to the nearest integer.

Proposition 2. *There exist large constants $C_0, c_1 > 0$ such that the following is true.*

Let $\log^{21} x \leq y \leq x^{1/100}$ be large, and let $(a_n)_{n \leq x}$ be any complex numbers with absolute values at most 1. Let $x^{-1/1000} < \delta \leq 1$, and let $0 \leq \theta_1, \theta_2, \dots, \theta_R \leq 1$ be any real numbers such that $\|\theta_r - \theta_s\| \geq 1/x$ when $r \neq s$, and such that

$$\left| \sum_{\substack{n \leq x, \\ n \in \mathcal{S}(y)}} a_n e(n\theta_r) \right| \geq \delta \Psi(x, y) \quad \forall 1 \leq r \leq R.$$

Then for any small $\epsilon > 0$ we have

$$R \ll_{\epsilon} \delta^{-2-20(1-\alpha(x,y))-\epsilon} \log^{3+32(1-\alpha(x,y))+\epsilon} x, \quad (4.1)$$

where $\alpha(x, y)$ denotes the saddle-point corresponding to the y -smooth numbers below x .

In addition, if $\log^{c_1} x \leq y \leq x^{1/100}$ and $\delta^{-1} \leq \min\{\log^{1/\epsilon} x, y^{1/C_0}\}$ then

$$R \ll_{\epsilon} \delta^{-2-8(1-\alpha(x,y))-\epsilon}. \quad (4.2)$$

First let us use Proposition 2 to prove Theorem 2. Recall that we have $1 - \alpha = \frac{\log(u \log(u+1)) + O(1)}{\log y} \leq 1/20$, in view of the approximation (2.1) and our assumption that $y \geq \log^{21} x$. If $\delta^{-1} > \log^{1/\epsilon} x$ then

$$\log^{3+32(1-\alpha)+\epsilon} x \leq \log^5 x \leq \delta^{-5\epsilon},$$

or if $y^{1/C_0} < \delta^{-1} \leq \log^{1/\epsilon} x$ then $u := (\log x)/\log y \gg_{\epsilon} (\log x)/\log \log x$, and therefore

$$\log^{3+32(1-\alpha)+\epsilon} x \leq \log^5 x \ll_{\epsilon} u^6 \ll y^{6(1-\alpha)} < \delta^{-6C_0(1-\alpha)}.$$

Thus we can bound the term $\log^{3+32(1-\alpha)+\epsilon} x$ in (4.1) by powers of δ in both of these cases, and the only remaining case is where $\delta^{-1} \leq \min\{\log^{1/\epsilon} x, y^{1/C_0}\}$, which is covered by (4.2). So relabelling the arbitrary small ϵ , we can combine (4.1) and (4.2) into the

single bound

$$R \ll_{\epsilon} \delta^{-2-(20+6C_0)(1-\alpha)-\epsilon}.$$

Applying this bound, together with Parseval's identity, we deduce that in Theorem 2 we have

$$\begin{aligned} & \int_0^1 \left| \sum_{\substack{n \leq x, \\ n \in \mathcal{S}(y)}} a_n e(n\theta) \right|^p d\theta \\ & \leq \sum_{1 \leq j \leq \frac{\log x}{1000 \log 2}} \left(\frac{\Psi(x, y)}{2^{j-1}} \right)^p \text{meas} \left\{ 0 \leq \theta \leq 1 : \frac{\Psi(x, y)}{2^j} < \left| \sum_{\substack{n \leq x, \\ n \in \mathcal{S}(y)}} a_n e(n\theta) \right| \leq \frac{\Psi(x, y)}{2^{j-1}} \right\} \\ & \quad + \left(\frac{\Psi(x, y)}{x^{1/1000}/2} \right)^{p-2} \int_0^1 \left| \sum_{\substack{n \leq x, \\ n \in \mathcal{S}(y)}} a_n e(n\theta) \right|^2 d\theta \\ & \ll_{p, \epsilon} \sum_{1 \leq j \leq \frac{\log x}{1000 \log 2}} \left(\frac{\Psi(x, y)}{2^j} \right)^p \cdot \frac{1}{x} 2^{j(2+(20+6C_0)(1-\alpha)+\epsilon)} + \frac{\Psi(x, y)^{p-1}}{x^{(p-2)/1000}}. \end{aligned}$$

We assume in Theorem 2 that $y \geq \log^{C \max\{1, 1/(p-2)\}} x$, which implies by (2.1) that

$$1 - \alpha(x, y) \leq \frac{\log \log x}{\log y} + o(1) \leq \frac{\min\{1, p-2\}}{C} + o(1),$$

and which also implies that $\Psi(x, y)/x \geq \Psi(x, \log^{C \max\{1, 1/(p-2)\}} x)/x \geq x^{-\min\{1, p-2\}/C + o(1)}$. So if we choose $\epsilon = (p-2)/3$, say, then provided C is large enough (so that in particular $(20 + 6C_0)(1 - \alpha) \leq (p-2)/3$) we will have

$$\sum_{1 \leq j \leq \frac{\log x}{1000 \log 2}} \left(\frac{\Psi(x, y)}{2^j} \right)^p \cdot \frac{1}{x} 2^{j(2+(20+6C_0)(1-\alpha)+\epsilon)} \leq \frac{\Psi(x, y)^p}{x} \sum_{1 \leq j \leq \frac{\log x}{1000 \log 2}} \frac{1}{2^{j(p-2)/3}} \ll_p \frac{\Psi(x, y)^p}{x},$$

and also $\Psi(x, y)^{p-1}/x^{(p-2)/1000} \ll \Psi(x, y)^p/x$. This is precisely Theorem 2.

Q.E.D.

It remains to prove Proposition 2, and we shall assume throughout, without loss of generality, that δ is at most a small absolute constant. We shall also concentrate on proving the bound (4.1). The bound (4.2), which is only relevant to save logarithmic factors when δ is fairly large (as seen in the combining of (4.1) and (4.2) in the above argument), can be proved by inserting Major Arc Estimates 1 and 2 into Bourgain's [3] original restriction argument, as we shall briefly explain at the end.

To try to make things more digestible, we split the proof of (4.1) into two halves. Firstly we shall prove the following intermediate statement, which already encodes much of the proof including the Bourgain/Halász–Montgomery duality step, our combination

of this with double sums, and the use of Theorem 1 and the Erdős–Turán inequality to reduce to looking at differences $\theta_r - \theta_s$ on “fairly major” arcs.

Proposition 3. *Let the situation be as in Proposition 2, with δ at most a small absolute constant, and define $K := \lfloor (\delta^{-1} \log x)^{10} \rfloor$ (where $\lfloor \cdot \rfloor$ denotes integer part).*

Then

$$\delta^2 R^2 \Psi(x, y) \ll \sum_{q \leq K \log^{12} x} \sum_{(a, q)=1} \sum_{\substack{1 \leq r, s \leq R, \\ \|\theta_r - \theta_s - \frac{a}{q}\| \leq \frac{K^2 y \log^{12} x}{x}}} \sum_{\substack{x/yK < m \leq x/K, \\ m \in \mathcal{S}(y)}} \min\left\{\frac{x}{m}, \frac{1}{\|m(\theta_r - \theta_s)\|}\right\}.$$

To prove Proposition 3, let $\theta_1, \dots, \theta_R$ be as in the statement, and let c_1, \dots, c_R be complex numbers of absolute value 1 such that $c_r \sum_{\substack{n \leq x, \\ n \in \mathcal{S}(y)}} a_n e(n\theta_r) = \left| \sum_{\substack{n \leq x, \\ n \in \mathcal{S}(y)}} a_n e(n\theta_r) \right|$. Then the Cauchy–Schwarz inequality yields that

$$\delta^2 R^2 \Psi(x, y)^2 \leq \left(\sum_{r=1}^R \left| \sum_{\substack{n \leq x, \\ n \in \mathcal{S}(y)}} a_n e(n\theta_r) \right| \right)^2 = \left(\sum_{r=1}^R c_r \sum_{\substack{n \leq x, \\ n \in \mathcal{S}(y)}} a_n e(n\theta_r) \right)^2 \leq \Psi(x, y) \sum_{\substack{n \leq x, \\ n \in \mathcal{S}(y)}} \left| \sum_{r=1}^R c_r e(n\theta_r) \right|^2.$$

This manipulation follows the first part of the argument in §4 of Bourgain [3], and multiplicative number theorists may be most familiar with it from the work of Halász and others on large values of Dirichlet polynomials.

At this point, Bourgain [3] proceeds by expanding the square, switching the order of summation, and inserting his good major arc estimates. But we do not have sufficient major arc information, so we shall proceed differently. Note that $1 < x/yK < \delta^3 x$, because of our assumptions that $y \leq x^{1/100}$ and $\delta > x^{-1/1000}$ (which implies that $K := \lfloor (\delta^{-1} \log x)^{10} \rfloor \leq x^{1/100+o(1)}$). Thus if we again write $P(m)$ for the largest prime factor of the integer m , and write $p(n)$ for the smallest prime factor of the integer n , then as we will explain we can rewrite the right hand sum as

$$\begin{aligned} \sum_{\substack{n \leq x, \\ n \in \mathcal{S}(y)}} \left| \sum_{r=1}^R c_r e(n\theta_r) \right|^2 &= \sum_{\substack{x/yK < m \leq x/K, \\ m/P(m) \leq x/yK, \\ m \in \mathcal{S}(y)}} \sum_{\substack{\delta^3 x/m \leq n \leq x/m, \\ p(n) \geq P(m), \\ n \in \mathcal{S}(y)}} \left| \sum_{r=1}^R c_r e(mn\theta_r) \right|^2 + \sum_{\substack{n < \delta^3 x, \\ n \in \mathcal{S}(y)}} \left| \sum_{r=1}^R c_r e(n\theta_r) \right|^2 \\ &\leq \sum_{\substack{x/yK < m \leq x/K, \\ m/P(m) \leq x/yK, \\ m \in \mathcal{S}(y)}} \sum_{\substack{\delta^3 x/m \leq n \leq x/m}} \left| \sum_{r=1}^R c_r e(mn\theta_r) \right|^2 + R^2 \Psi(\delta^3 x, y) \\ &\ll \sum_{\substack{x/yK < m \leq x/K, \\ m/P(m) \leq x/yK, \\ m \in \mathcal{S}(y)}} \sum_{\substack{1 \leq r, s \leq R}} \min\left\{\frac{x}{m}, \frac{1}{\|m(\theta_r - \theta_s)\|}\right\} + R^2 \Psi(x, y) \delta^{3\alpha}. \end{aligned}$$

Indeed the first equality follows by decomposing each number $\geq \delta^3 x$ in the sum on the left into two factors m, n containing small and large prime factors (similarly as at the start of the proof of Theorem 1), and the final inequality follows by expanding the square and performing the summation over n , and by using Smooth Numbers Result 2. In particular, since $y \geq \log^{21} x$ we have $\alpha(x, y) \geq 19/20$, by the approximation (2.1). So provided δ is sufficiently small the contribution from the final term will certainly be $\leq \delta^2 R^2 \Psi(x, y)/2$, and since overall we have at least $\delta^2 R^2 \Psi(x, y)$ on the left hand side we must have

$$\delta^2 R^2 \Psi(x, y) \ll \sum_{\substack{x/yK < m \leq x/K, \\ m \in \mathcal{S}(y)}} \sum_{1 \leq r, s \leq R} \min\left\{\frac{x}{m}, \frac{1}{\|m(\theta_r - \theta_s)\|}\right\}.$$

To finish the proof of Proposition 3, we need to show that the only pairs r, s that can make a non-negligible contribution to the right hand side are those where $\theta_r - \theta_s$ is fairly close to a rational with fairly small denominator (in terms of $K := \lfloor (\delta^{-1} \log x)^{10} \rfloor$). One could presumably show this directly by combinatorial arguments, but this would entail a further careful decomposition of the m variable. Instead after a little bit of preparation we shall use the Erdős–Turán inequality (Harmonic Analysis Result 1), together with Theorem 1. Indeed, for any $r, s \leq R$ we have

$$\sum_{\substack{\frac{x}{yK} < m \leq \frac{x}{K}, \\ m \in \mathcal{S}(y)}} \min\left\{\frac{x}{m}, \frac{1}{\|m(\theta_r - \theta_s)\|}\right\} \leq \Psi\left(\frac{x}{K}, y\right) \sqrt{K} + \sum_{\substack{x/yK < m \leq x/K, \\ m \in \mathcal{S}(y), \\ \|m(\theta_r - \theta_s)\| \leq 1/\sqrt{K}}} \min\left\{\frac{x}{m}, \frac{1}{\|m(\theta_r - \theta_s)\|}\right\},$$

and using Smooth Numbers Result 2 it follows that

$$\sum_{\substack{\frac{x}{yK} < m \leq \frac{x}{K}, \\ m \in \mathcal{S}(y)}} \min\left\{\frac{x}{m}, \frac{1}{\|m(\theta_r - \theta_s)\|}\right\} \ll \Psi(x, y) K^{1/2-\alpha} + x \sum_{\substack{x/yK < m \leq x/K, \\ m \in \mathcal{S}(y), \\ \|m(\theta_r - \theta_s)\| \leq 1/\sqrt{K}}} \frac{1}{m}.$$

For simplicity of writing, for any $x/yK < M \leq x/K$ let us temporarily write $\sum_{m \sim M}$ to mean $\sum_{\substack{M < m \leq 2M, \\ m \in \mathcal{S}(y), \\ \|m(\theta_r - \theta_s)\| \leq 1/\sqrt{K}}}$. Then the Erdős–Turán inequality implies that, for any $J \in \mathbb{N}$,

$$\sum_{m \sim M} 1 \ll \frac{\Psi(2M, y)}{\sqrt{K}} + \frac{\Psi(2M, y)}{J+1} + \sum_{j=1}^J \frac{1}{j} \left| \sum_{\substack{M < m \leq 2M, \\ m \in \mathcal{S}(y)}} e(jm(\theta_r - \theta_s)) \right|.$$

In particular, if we choose $J := \lfloor \sqrt{K} \rfloor$, and if $\theta_r - \theta_s = a/q + \eta$ for some $q \leq x^{0.6}$, some $(a, q) = 1$ and some $|\eta| \leq 1/qx^{0.6}$, say, then

$$\sum_{m \sim M} 1 \ll \frac{\Psi(2M, y)}{\sqrt{K}} + \sum_{h|q} \frac{1}{h} \sum_{\substack{j' \leq \sqrt{K}/h, \\ (j', q/h)=1}} \frac{1}{j'} \left| \sum_{\substack{M < m \leq 2M, \\ m \in \mathcal{S}(y)}} e(mj'h(\theta_r - \theta_s)) \right|.$$

In the exponential sum here we will have $j'h(\theta_r - \theta_s) = aj'/(q/h) + j'h\eta$, where aj' is coprime to q/h and where $|j'h\eta| \leq \sqrt{K}/qx^{0.6} \leq 1/qx^{0.55}$.

Now in the case where $q/h \leq x^{0.45}$, the condition $(q/h)^2 y^3 (1 + |j'h\eta M|)^2 \leq M/4$ in Theorem 1 will be satisfied, remembering that $y \leq x^{1/100}$ throughout this section, that $|j'h\eta| \leq 1/qx^{0.55}$, and that $M > x/yK \geq x^{0.98+o(1)}$. Then the theorem implies that

$$\frac{1}{h} \left| \sum_{\substack{M < m \leq 2M, \\ m \in \mathcal{S}(y)}} e(mj'h(\theta_r - \theta_s)) \right| \ll \frac{1}{h} \frac{\Psi(2M, y) \log^{7/2} x}{((q/h)(1 + |j'h\eta M|))^{1/2-3/2(1-\alpha)}} \ll \frac{\Psi(2M, y) \log^{7/2} x}{(q(1 + |\eta M|))^{1/2-3/2(1-\alpha)}}.$$

If $x^{0.45} < q/h \leq x^{0.6}$ then we can instead apply Fouvry and Tenenbaum's [13] minor arc estimate, as stated in our survey in the Introduction (see (1.1)), obtaining a bound $\ll M/x^{0.15}$ (say) that is negligible compared with the term $\Psi(2M, y)/\sqrt{K}$. We deduce overall that, since $1 - \alpha \leq 1/20$ (by the approximation (2.1) and our assumption that $y \geq \log^{21} x$),

$$\sum_{m \sim M} 1 \ll \frac{\Psi(2M, y)}{\sqrt{K}} + \frac{\Psi(2M, y) d(q) \log^{9/2} x}{(q(1 + |\eta M|))^{1/2-3/40}} \ll \frac{\Psi(2M, y)}{\sqrt{K}} + \frac{\Psi(2M, y) \log^{9/2} x}{(q(1 + |\eta x/yK|))^{2/5}},$$

where $d(q) = \sum_{h|q} 1$ denotes the divisor function. Here the second inequality used the facts that $d(q) = q^{o(1)}$ and $M > x/yK$.

Finally, we note in particular that if either $q \geq K \log^{12} x$ or $|\eta| \geq K^2 y (\log^{12} x)/x$ then the above is $\ll \Psi(2M, y)/K^{2/5}$. By using Smooth Numbers Result 2 in the form $\Psi(2M, y)/M \ll (\Psi(x, y)/x)(x/M)^{1-\alpha} \leq (\Psi(x, y)/x)(yK)^{1-\alpha}$, and summing over logarithmically many dyadic values of M , it follows that

$$x \sum_{\substack{x/yK < m \leq x/K, \\ m \in \mathcal{S}(y), \\ \|m(\theta_r - \theta_s)\| \leq 1/\sqrt{K}}} \frac{1}{m} \ll \frac{\Psi(x, y)}{K^{2/5}} \log y (yK)^{1-\alpha} = \Psi(x, y) K^{3/5-\alpha} y^{1-\alpha} \log y.$$

Then we have $K^{3/5-\alpha} \leq K^{3/5-19/20} = K^{-7/20} \ll \delta^{7/2}/\log^{7/2} x$, by definition of $K := \lfloor (\delta^{-1} \log x)^{10} \rfloor$. We also have $y^{1-\alpha} \log y \ll u \log(u+1) \log y \ll \log^2 x$ by the approximation (2.1). We conclude that if either $q \geq K \log^{12} x$ or $|\eta| \geq K^2 y (\log^{12} x)/x$ then the contribution from $\theta_r - \theta_s$ is $\ll \Psi(x, y) \delta^{7/2}/\log^{3/2} x$, which is indeed negligible compared with $\delta^2 \Psi(x, y)$ and may be discarded. It follows that we may restrict attention to those pairs of $1 \leq r, s \leq R$ for which $\|\theta_r - \theta_s - \frac{a}{q}\| \leq K^2 y (\log^{12} x)/x$ for some $q \leq K \log^{12} x$. Proposition 3 follows immediately.

Q.E.D.

To continue the proof of (4.1), we need to bound $\sum_{\substack{x/yK < m \leq x/K, \\ m \in \mathcal{S}(y)}} \min\left\{\frac{x}{m}, \frac{1}{\|m(\theta_r - \theta_s)\|}\right\}$ when $\|\theta_r - \theta_s - \frac{a}{q}\| \leq \frac{K^2 y \log^{12} x}{x}$ for some $q \leq K \log^{12} x$, as in the conclusion of Proposition 3. In fact we shall prove the following result.

Proposition 4. *Let the situation be as in Proposition 3. Suppose that $q \leq K \log^{12} x$, that $(a, q) = 1$, and that $|\eta| \leq K^2 y (\log^{12} x)/x$. Also define $\widetilde{M} := \min\{1/2q|\eta|, x/K\}$.*

Then if $\theta = a/q + \eta$, we have the bounds

$$\sum_{\substack{\widetilde{M} < m \leq x/K, \\ m \in \mathcal{S}(y)}} \min\left\{\frac{x}{m}, \frac{1}{\|m\theta\|}\right\} \ll \frac{\delta^2 \Psi(x, y)}{\log x},$$

$$\text{and} \quad \sum_{\substack{x/yK < m \leq \widetilde{M}, \\ m \in \mathcal{S}(y)}} \min\left\{\frac{x}{m}, \frac{1}{\|m\theta\|}\right\} \ll \frac{\delta^2 \Psi(x, y)}{\log x} + \frac{\Psi(x, y)}{q(1 + |\eta x|)} \log^2 x (K^2 \log^{12} x)^{1-\alpha(x, y)}.$$

The proof of Proposition 4 is a counting argument depending on Smooth Numbers Results 2 and 3, similarly as the proof of Theorem 1.

We begin with the first bound in Proposition 4, which is the harder one. We may assume that $|\eta| \geq K/2qx$ and therefore $\widetilde{M} = 1/2q|\eta|$, since otherwise the range of summation in the bound is empty and the statement is trivial. We may also assume without loss of generality (on negating θ) that $\eta > 0$. Observe then that if m lies in an interval of the form $(j/qK\eta, (j+1)/qK\eta]$ then we have $|m\theta - ma/q - j/qK| \leq 1/qK$. Further, if $j \equiv k$ modulo K , where we choose the representative k such that $-K/2 < k \leq K/2$, then for a certain integer $f(j)$ we will have

$$|m\theta - ma/q - f(j)/q - k/qK| \leq 1/qK.$$

Therefore as m ranges over the interval and over all residue classes modulo q , except for the unique class such that $ma \equiv -f(j)$ modulo q , the numbers $\|m\theta\| \asymp \|ma/q + f(j)/q\|$ will be $\asymp b/q$ for those $1 \leq b \leq q/2$. Meanwhile, if $|k| \geq 2$ then, when m is in the interval and $ma \equiv -f(j)$ modulo q , we will have $\|m\theta\| \asymp |k|/qK$. It is only if $ma \equiv -f(j)$ modulo q and $k \in \{-1, 0, 1\}$ that we cannot usefully lower bound $\|m\theta\|$.

In view of the above discussion, if we split the sum over m into intervals $(j/qK\eta, (j+1)/qK\eta]$ (for varying j) and further into arithmetic progressions $b \bmod q$, we see

$$\begin{aligned} & \sum_{\substack{\widetilde{M} < m \leq x/K, \\ m \in \mathcal{S}(y)}} \min\left\{\frac{x}{m}, \frac{1}{\|m\theta\|}\right\} \leq \sum_{-K/2 < k \leq K/2} \sum_{\substack{qK\widetilde{M}\eta - 1 \leq j \leq qx\eta, \\ j \equiv k \bmod K}} \sum_{b=1}^q \sum_{\substack{j/qK\eta < m \leq (j+1)/qK\eta, \\ m \equiv b \bmod q, \\ m \in \mathcal{S}(y)}} \min\left\{\frac{x}{m}, \frac{1}{\|m\theta\|}\right\} \\ & \ll \sum_{|k| \leq \frac{K}{2}} \sum_{\substack{j \leq qx\eta, \\ j \equiv k \bmod K}} \left(\sum_{b=1}^q \frac{q}{b} \right) \max_{1 \leq b \leq q} \sum_{\substack{j/qK\eta < m \leq (j+1)/qK\eta, \\ m \equiv b \bmod q, \\ m \in \mathcal{S}(y)}} 1 + \sum_{\substack{|k| \leq \frac{K}{2}, \\ k \notin \{-1, 0, 1\}}} \frac{qK}{|k|} \sum_{\substack{j \leq qx\eta, \\ j \equiv k \bmod K}} \sum_{\substack{j/qK\eta < m \leq (j+1)/qK\eta, \\ m \equiv -a^{-1}f(j) \bmod q, \\ m \in \mathcal{S}(y)}} 1 \\ & + \sum_{-1 \leq k \leq 1} \sum_{\substack{qK\widetilde{M}\eta - 1 \leq j \leq qx\eta, \\ j \equiv k \bmod K}} \sum_{\substack{j/qK\eta < m \leq (j+1)/qK\eta, \\ m \equiv -a^{-1}f(j) \bmod q, \\ m \in \mathcal{S}(y)}} \frac{x}{m}. \end{aligned} \tag{4.3}$$

Each of the first two terms in (4.3) is

$$\begin{aligned} &\ll qK \log(q+1) \sum_{|k| \leq \frac{K}{2}} \frac{1}{|k|+1} \sum_{\substack{j \leq qx\eta, \\ j \equiv k \pmod{K}}} \max_{1 \leq b \leq q} \sum_{\substack{\frac{j}{qK\eta} < m \leq \frac{j+1}{qK\eta}, \\ m \equiv b \pmod{q}, \\ m \in \mathcal{S}(y)}} 1 \\ &\ll qK \log(q+1) \sum_{|k| \leq \frac{K}{2}} \frac{1}{|k|+1} \sum_{\substack{j \leq qx\eta, \\ j \equiv k \pmod{K}}} \frac{\Psi(j/qK\eta, y)}{(jq)^\alpha} \log x, \end{aligned}$$

by Smooth Numbers Result 3. Further, using Smooth Numbers Result 2 to bound $\Psi(j/qK\eta, y) \ll \Psi(x, y)(j/qK\eta)^\alpha$ we find the above is

$$\ll qK \log(q+1) \sum_{|k| \leq \frac{K}{2}} \frac{1}{|k|+1} \sum_{\substack{j \leq qx\eta, \\ j \equiv k \pmod{K}}} \frac{\Psi(x, y)}{q^2 K \eta x} (q^2 K \eta x)^{1-\alpha} \log x.$$

Since we assume that $\eta = |\eta| \geq K/2qx$ we have $qx\eta \geq K/2$, and so the above is

$$\ll qK \log(q+1) \sum_{|k| \leq \frac{K}{2}} \frac{1}{|k|+1} \frac{\Psi(x, y)}{qK^2} (q^2 K \eta x)^{1-\alpha} \log x \ll \frac{\Psi(x, y)}{K} \log^3 x (q^2 K \eta x)^{1-\alpha}.$$

Similarly, using the bound $x/m < xqK\eta/j$ the third term in (4.3) is seen to be

$$\ll xqK\eta \sum_{-1 \leq k \leq 1} \sum_{\substack{qK\widetilde{M}\eta-1 \leq j \leq qx\eta, \\ j \equiv k \pmod{K}}} \frac{1}{j} \max_{1 \leq b \leq q} \sum_{\substack{\frac{j}{qK\eta} < m \leq \frac{j+1}{qK\eta}, \\ m \equiv b \pmod{q}, \\ m \in \mathcal{S}(y)}} 1,$$

and using the Smooth Numbers Results as before this is

$$\ll \frac{\Psi(x, y)}{q} (q^2 K \eta x)^{1-\alpha} \log x \sum_{-1 \leq k \leq 1} \sum_{\substack{qK\widetilde{M}\eta-1 \leq j \leq qx\eta, \\ j \equiv k \pmod{K}}} \frac{1}{j} \ll \frac{\Psi(x, y)}{qK} \log^2 x (q^2 K \eta x)^{1-\alpha}.$$

Here the final inequality crucially uses the fact that $qK\widetilde{M}\eta = K/2$, and therefore the smallest terms in the sum over j are $\asymp K$, and the sum itself is $\ll (\log x)/K$.

Since $q^2 K \eta x \leq q^2 K^3 y \log^{12} x \leq K^5 y \log^{36} x$, both of the above bounds are

$$\ll \frac{\Psi(x, y)}{K} \log^3 x (K^5 \log^{36} x)^{1-\alpha} y^{1-\alpha}.$$

Moreover, we have $y^{1-\alpha} \ll u \log(u+1) \ll \log x$ and $1-\alpha \leq 1/20$ by the approximation (2.1) (and the fact that $y \geq \log^{21} x$), so the bounds are

$$\ll \frac{\Psi(x, y)}{K^{3/4}} \log^{29/5} x.$$

Since $K \gg (\delta^{-1} \log x)^{10}$ this is all $\ll \Psi(x, y) \delta^{30/4} / \log^{17/10} x \ll \Psi(x, y) \delta^2 / \log x$, which is precisely the first bound claimed in Proposition 4.

It remains to prove the second bound in Proposition 4, for $\sum_{\substack{x/yK < m \leq \widetilde{M}, \\ m \in \mathcal{S}(y)}} \min\{\frac{x}{m}, \frac{1}{|m\theta|}\}$.

This time we don't need to split the range of m into intervals, since we have $|m\eta| \leq \widetilde{M}|\eta| \leq 1/2q$ by definition of \widetilde{M} , and therefore we simply have

$$|m\theta| = \left| \frac{ma}{q} + m\eta \right| \asymp \begin{cases} \left| \frac{ba}{q} \right| & \text{if } m \equiv b \pmod{q}, \text{ for } b \neq 0 \\ |m\eta| & \text{if } m \equiv 0 \pmod{q}. \end{cases}$$

So we find, similarly as above (and remembering that $\widetilde{M} \leq x/K$), that

$$\begin{aligned} \sum_{\substack{x/yK < m \leq \widetilde{M}, \\ m \in \mathcal{S}(y)}} \min\left\{\frac{x}{m}, \frac{1}{|m\theta|}\right\} &\ll \left(\sum_{b=1}^{q-1} \frac{q}{b}\right) \max_{1 \leq b \leq q-1} \sum_{\substack{x/yK < m \leq \widetilde{M}, \\ m \equiv b \pmod{q}, \\ m \in \mathcal{S}(y)}} 1 + \sum_{\substack{x/yK < m \leq \widetilde{M}, \\ m \equiv 0 \pmod{q}, \\ m \in \mathcal{S}(y)}} \min\left\{\frac{x}{m}, \frac{1}{m|\eta|}\right\} \\ &\ll q \log q \frac{\Psi(\frac{x}{K}, y)}{q} q^{1-\alpha} \log x + \frac{x}{q} \min\left\{1, \frac{1}{|\eta x|}\right\} \sum_{\substack{x/qyK < m' \leq x/qK, \\ m' \in \mathcal{S}(y)}} \frac{1}{m'}. \end{aligned}$$

Then Smooth Numbers Result 2 and the fact that $\alpha \geq 19/20$ imply that $\Psi(x/K, y) \ll \Psi(x, y)/K^\alpha \ll \Psi(x, y)/K^{19/20}$, and so the first term is $\ll \Psi(x, y)q^{1/20}(\log^2 x)/K^{19/20}$. Since we also have $q \leq K \log^{12} x$ the first term is $\ll \Psi(x, y)(\log^{13/5} x)/K^{9/10}$, and since $K := \lfloor (\delta^{-1} \log x)^{10} \rfloor$ this is certainly all $\ll \Psi(x, y)\delta^2/\log x$. Meanwhile, by dividing into dyadic intervals and using Smooth Numbers Result 2 we find

$$\sum_{\substack{x/qyK < m' \leq x/qK, \\ m' \in \mathcal{S}(y)}} \frac{1}{m'} \ll \log y \max_{x/qyK < M \leq x/qK} \frac{\Psi(M, y)}{M} \ll \frac{\Psi(x, y)}{x} \log y (qyK)^{1-\alpha},$$

and by using again the fact that $y^{1-\alpha} \ll u \log(u+1) = (\log x \log(u+1))/\log y$ we find the sum is

$$\ll \frac{\Psi(x, y)}{x} \log^2 x (qK)^{1-\alpha}.$$

Since $q \leq K \log^{12} x$ we have $(qK)^{1-\alpha} \leq (K^2 \log^{12} x)^{1-\alpha}$, and we finally conclude that

$$\sum_{\substack{x/yK < m \leq \widetilde{M}, \\ m \in \mathcal{S}(y)}} \min\left\{\frac{x}{m}, \frac{1}{|m\theta|}\right\} \ll \frac{\delta^2 \Psi(x, y)}{\log x} + \frac{\Psi(x, y)}{q(1+|\eta x|)} \log^2 x (K^2 \log^{12} x)^{1-\alpha},$$

as claimed in Proposition 4.

Q.E.D.

Finally, by combining Propositions 3 and 4 we obtain that

$$\delta^2 R^2 \Psi(x, y) \ll \sum_{q \leq K \log^{12} x} \sum_{(a, q)=1} \sum_{1 \leq r, s \leq R} \frac{\Psi(x, y)}{q(1+x|\theta_r - \theta_s - a/q|)} \log^2 x (K^2 \log^{12} x)^{1-\alpha}.$$

Here we noted that for each pair of $1 \leq r, s \leq R$ there is at most one pair a, q such that $|\theta_r - \theta_s - \frac{a}{q}| \leq \frac{K^2 y \log^{12} x}{x}$, given that $y \leq x^{1/100}$ and $K \leq x^{1/100+o(1)}$. (This means that

the contribution from all terms $\frac{\delta^2 \Psi(x, y)}{\log x}$ in Proposition 4 to Proposition 3 is $\ll \frac{\delta^2 R^2 \Psi(x, y)}{\log x}$, which is negligible and may be discarded).

If we now apply Harmonic Analysis Result 2, with the choices $Q = K \log^{12} x$ and $v = 1/2$ and $A = 1$, we deduce that

$$\delta^2 R^2 \ll_{\epsilon} \log^2 x (K^2 \log^{12} x)^{1-\alpha} \left(R Q^{\epsilon} \log x + \frac{R^2 Q \log x}{x} + \frac{R^2 \log x}{Q} \right).$$

Since $1 - \alpha \leq 1/20$ and $K = \lfloor (\delta^{-1} \log x)^{10} \rfloor$ the contribution from the last term is negligible compared with $\delta^2 R^2$. Since $\delta \geq x^{-1/1000}$ and therefore $Q \leq x^{1/100+o(1)}$ the second term is also negligible. Discarding them and rearranging we obtain

$$R \ll_{\epsilon} \log^3 x (K^2 \log^{12} x)^{1-\alpha} \delta^{-2} Q^{\epsilon} \ll_{\epsilon} \delta^{-2-20(1-\alpha)-10\epsilon} \log^{3+32(1-\alpha)+22\epsilon} x,$$

which is precisely the claimed bound (4.1) after relabelling the arbitrary small ϵ .

Q.E.D.

As promised, we end this section by briefly sketching how to obtain the other bound (4.2). Recall that we now assume $\delta^{-1} \leq \min\{\log^{1/\epsilon} x, y^{1/C_0}\}$. Proceeding as in §4 of Bourgain [3], or as at the very beginning of the proof of Proposition 3, we obtain that

$$\delta^2 R^2 \Psi(x, y) \leq \sum_{1 \leq r, s \leq R} \left| \sum_{\substack{n \leq x, \\ n \in \mathcal{S}(y)}} e(n(\theta_r - \theta_s)) \right|.$$

Now combining Theorem 1 with Dirichlet's approximation theorem, as previously, the only way that we can possibly have $\left| \sum_{\substack{n \leq x, \\ n \in \mathcal{S}(y)}} e(n(\theta_r - \theta_s)) \right| \geq (\delta^2/2) \Psi(x, y)$ is if $\theta_r - \theta_s = a/q + \eta$ for some $q \leq (\delta^{-1} \log x)^{10}$ and some $|\eta| \leq (\delta^{-1} \log x)^{10}/x$, say. Moreover, since we have $\delta^{-1} \leq \min\{\log^{1/\epsilon} x, y^{1/C_0}\}$ we can apply Major Arc Estimates 1 and 2 in that case, and deduce the stronger fact that actually

$$\begin{aligned} \sum_{\substack{n \leq x, \\ n \in \mathcal{S}(y)}} e(n(\frac{a}{q} + \eta)) &= V(x, y; q, \eta) + O_{\epsilon} \left(\frac{\delta^2 \Psi(x, y)}{\log x} \right) \\ &\ll_{\epsilon} \Psi(x, y) \frac{2^{\omega(q)} q^{1-\alpha} \log^2(q+1) \log^3(2+|\eta x|)}{\phi(q)} \frac{1}{(1+|\eta x|)^{\alpha}} + \frac{\delta^2 \Psi(x, y)}{\log x}. \end{aligned}$$

Here we bounded the main term $(\Psi(x, y)/x) V(x, y; q, \delta)$ in Major Arc Estimate 2 by noting it is $(\Psi(x, y)/x) \sum_{n \leq x} e(n(a/q + \eta)) + O_{\epsilon}(\delta^2 \Psi(x, y)/\log x)$, (by Major Arc Estimate 1 again), and then summing the geometric progression.

Using the stronger bound, provided δ is small enough in terms of ϵ the only way that we can have $\left| \sum_{\substack{n \leq x, \\ n \in \mathcal{S}(y)}} e(n(\theta_r - \theta_s)) \right| \geq (\delta^2/2) \Psi(x, y)$ is actually if $\theta_r - \theta_s = a/q + \eta$ for some $q \leq \delta^{-4}$ and some $|\eta| \leq \delta^{-4}/x$, say. In that case, on noting that $\log(q +$

1), $2^{\omega(q)}$, $q/\phi(q)$ are all $\ll_{\epsilon} q^{\epsilon/100} \leq \delta^{-\epsilon/25}$ (say), similarly for $\log(2 + |\eta x|)$, we will have

$$\left| \sum_{\substack{n \leq x, \\ n \in \mathcal{S}(y)}} e(n(\theta_r - \theta_s)) \right| \ll_{\epsilon} \Psi(x, y) \frac{q^{1-\alpha} \delta^{-\epsilon}}{q(1 + |\eta x|)^{\alpha}} + \frac{\delta^2 \Psi(x, y)}{\log x} \ll_{\epsilon} \Psi(x, y) \frac{\delta^{-8(1-\alpha)-\epsilon}}{q(1 + |\eta x|)} + \frac{\delta^2 \Psi(x, y)}{\log x}.$$

So overall, on discarding all terms that are $< (\delta^2/2)\Psi(x, y)$, we must have

$$\delta^2 R^2 \Psi(x, y) \ll_{\epsilon} \delta^{-8(1-\alpha)-\epsilon} \sum_{q \leq \delta^{-4}} \sum_{(a, q)=1} \sum_{\substack{1 \leq r, s \leq R, \\ \|\theta_r - \theta_s - a/q\| \leq \delta^{-4}/x}} \frac{\Psi(x, y)}{q(1 + x\|\theta_r - \theta_s - a/q\|)}.$$

Notice there are no wasteful logarithmic factors on the right hand side of this inequality.

The bound (4.2) now follows swiftly as at the end of the proof of (4.1), by applying Harmonic Analysis Result 2 to the right hand side with the choices $Q = \delta^{-4}$, $v = \delta^{-4}/x$ and $A = 1$.

Q.E.D.

5. PROOF OF COROLLARY 1

The proof of Corollary 1 will be an application of the circle method, using Theorem 2, Theorem 1 and the major arc estimates in §2.2. Set $R := \log^{20} x$, and

$$\mathfrak{M} := \bigcup_{q \leq R} \bigcup_{(a, q)=1} [a/q - R/x, a/q + R/x].$$

In order that we may apply Theorem 1 straightforwardly, we assume throughout this section that $\log^K x \leq y \leq x^{1/100}$. The case of Corollary 1 where $x^{1/100} < y \leq x$ is handled by the work of La Bretèche and Granville [6], for example.

By the orthogonality of additive characters,

$$\#\{(a, b, c) \in \mathcal{S}(y)^3 : a, b, c \leq x, a + b = c\} = \int_0^1 \left(\sum_{\substack{n \leq x, \\ n \in \mathcal{S}(y)}} e(n\theta) \right)^2 \overline{\left(\sum_{\substack{n \leq x, \\ n \in \mathcal{S}(y)}} e(n\theta) \right)} d\theta.$$

The main term $\Psi(x, y)^3/2x$ in Corollary 1 comes from the integral over \mathfrak{M} , and the error term $O(\Psi(x, y)^3 \log(u+1)/(x \log y))$ comes partly from \mathfrak{M} and also from the rest of the integral. To extract the main term neatly, let us also note that

$$\frac{x^3}{2x} (1 + O(1/x)) = \#\{(a, b, c) : a, b, c \leq x, a + b = c\} = \int_0^1 \left(\sum_{n \leq x} e(n\theta) \right)^2 \overline{\left(\sum_{n \leq x} e(n\theta) \right)} d\theta,$$

by a trivial counting argument. We will show that

$$\int_{[0,1] \setminus \mathfrak{M}} \left| \sum_{\substack{n \leq x, \\ n \in \mathcal{S}(y)}} e(n\theta) \right|^3 d\theta \ll \frac{\Psi(x, y)^3}{x} \frac{1}{\log^4 x}, \quad \text{and} \quad \int_{[0,1] \setminus \mathfrak{M}} \left| \sum_{n \leq x} e(n\theta) \right|^3 d\theta \ll \frac{x^3}{x} \frac{1}{\log^4 x},$$

and we will also show that

$$\begin{aligned} \int_{\mathfrak{M}} \left(\sum_{\substack{n \leq x, \\ n \in \mathcal{S}(y)}} e(n\theta) \right)^2 \overline{\left(\sum_{\substack{n \leq x, \\ n \in \mathcal{S}(y)}} e(n\theta) \right)} d\theta &= \frac{\Psi(x, y)^3}{x^3} \int_{\mathfrak{M}} \left(\sum_{n \leq x} e(n\theta) \right)^2 \overline{\left(\sum_{n \leq x} e(n\theta) \right)} d\theta \\ &\quad + O\left(\frac{\Psi(x, y)^3 \log(u+1)}{x \log y} \right). \end{aligned} \quad (5.1)$$

Combining all these facts immediately yields the corollary.

Before the main proof, we shall demonstrate the following proposition.

Proposition 5. *For any large $\log^K x \leq y \leq x^{1/100}$, and any $\theta \in [0, 1] \setminus \mathfrak{M}$, we have*

$$\left| \sum_{\substack{n \leq x, \\ n \in \mathcal{S}(y)}} e(n\theta) \right| \ll \Psi(x, y) \frac{1}{\log^5 x}.$$

To prove the proposition we distinguish two different cases. By Dirichlet's approximation theorem, for each $\theta \in [0, 1]$ we have $\theta = a/q + \delta$ for some $q \leq x^{0.55}$, some $(a, q) = 1$, and some $|\delta| \leq 1/qx^{0.55}$, and if $\theta \in [0, 1] \setminus \mathfrak{M}$ then we either have $q > R$ or $|\delta| > R/x$. If $\log^K x \leq y \leq x^{1/100}$ and $x^{0.48} \leq q \leq x^{0.55}$ then, by Fouvry and Tenenbaum's [13] minor arc estimate stated in the introduction,

$$\left| \sum_{\substack{n \leq x, \\ n \in \mathcal{S}(y)}} e(n\theta) \right| \ll x(1+|\delta x|) \log^3 x \left(\frac{\sqrt{y}}{x^{1/4}} + \frac{1}{\sqrt{q}} + \sqrt{\frac{qy}{x}} \right) \ll x \log^3 x \left(\frac{1}{x^{0.245}} + \frac{1}{x^{0.24}} + \frac{1}{x^{0.22}} \right),$$

on noting that $|\delta x| \leq x^{0.45}/q \leq 1$. Since $\Psi(x, y) \geq \Psi(x, \log^K x) = x^{1-1/K+o(1)}$, this estimate is much stronger than we need provided K is large enough. The other case is where $\log^K x \leq y \leq x^{1/100}$, $q < x^{0.48}$, and either $q > R := \log^{20} x$ or $|\delta| > (\log^{20} x)/x$. Then we check that $q^2 y^3 (1+|\delta x|)^2 \leq x/4$, and we have $1 - \alpha(x, y) \leq 1/K + o(1) \leq 1/100$ (say) by the approximation (2.1), so Theorem 1 is applicable and yields that

$$\left| \sum_{\substack{n \leq x, \\ n \in \mathcal{S}(y)}} e(n\theta) \right| \ll \frac{\Psi(x, y)}{(\log^{20} x)^{1/2-(3/2)(1-\alpha)}} \log^{7/2} x \ll \frac{\Psi(x, y)}{(\log^{20} x)^{0.45}} \log^{7/2} x \ll \frac{\Psi(x, y)}{\log^5 x}.$$

Q.E.D.

Now we can deduce our claimed bounds for the integrals over $[0, 1] \setminus \mathfrak{M}$. Provided $y \geq \log^K x$ with K large enough, Theorem 2 implies that

$$\int_{[0, 1] \setminus \mathfrak{M}} \left| \sum_{\substack{n \leq x, \\ n \in \mathcal{S}(y)}} e(n\theta) \right|^3 \leq \sup_{\theta \notin \mathfrak{M}} \left| \sum_{\substack{n \leq x, \\ n \in \mathcal{S}(y)}} e(n\theta) \right|^{0.9} \int_0^1 \left| \sum_{\substack{n \leq x, \\ n \in \mathcal{S}(y)}} e(n\theta) \right|^{2.1} \ll \frac{\Psi(x, y)^{2.1}}{x} \sup_{\theta \notin \mathfrak{M}} \left| \sum_{\substack{n \leq x, \\ n \in \mathcal{S}(y)}} e(n\theta) \right|^{0.9}.$$

Then Proposition 5 implies the above is all $\ll (\Psi(x, y)^3/x)(1/\log^4 x)$, as claimed. The bound $\int_{[0,1] \setminus \mathfrak{M}} |\sum_{n \leq x} e(n\theta)|^3 d\theta \ll \frac{x^3}{x} \frac{1}{\log^4 x}$ for the complete sum is a trivial consequence of the pointwise bound $\sum_{n \leq x} e(n\theta) \ll 1/||\theta||$.

It only remains to show the approximate scaling relation (5.1) between the integrals over \mathfrak{M} . But when $\theta = a/q + \delta \in \mathfrak{M}$ (so that $q, |\delta x| \leq R := \log^{20} x$), and provided $y \geq \log^K x$ with K large enough, Major Arc Estimates 1 and 2 imply that

$$\begin{aligned} \sum_{\substack{n \leq x, \\ n \in \mathcal{S}(y)}} e(n\theta) &= \frac{\Psi(x, y)}{x} V(x, x; q, \delta) + O\left(\Psi(x, y) \frac{\log(u+1)}{\log y} \frac{2^{\omega(q)} q^{1-\alpha} \log^2(q+1) \log^3(2+|\delta x|)}{\phi(q) (1+|\delta x|)^\alpha}\right) \\ &= \frac{\Psi(x, y)}{x} \sum_{n \leq x} e(n\theta) + O\left(\Psi(x, y) \frac{\log(u+1)}{\log y} \frac{2^{\omega(q)} q^{1-\alpha} \log^2(q+1) \log^3(2+|\delta x|)}{\phi(q) (1+|\delta x|)^\alpha}\right). \end{aligned}$$

Here the second equality follows by using Major Arc Estimate 1 again, to compare $V(x, x; q, \delta)$ with $\sum_{n \leq x} e(n\theta)$. Moreover, we certainly always have $\sum_{n \leq x} e(n\theta) \ll \min\{x, 1/||\theta||\} \ll x \frac{2^{\omega(q)} q^{1-\alpha} \log^2(q+1) \log^3(2+|\delta x|)}{\phi(q) (1+|\delta x|)^\alpha}$ when $\theta \in \mathfrak{M}$, and therefore

$$\begin{aligned} &\left| \int_{\mathfrak{M}} \left(\sum_{\substack{n \leq x, \\ n \in \mathcal{S}(y)}} e(n\theta) \right)^2 \overline{\left(\sum_{\substack{n \leq x, \\ n \in \mathcal{S}(y)}} e(n\theta) \right)} d\theta - \frac{\Psi(x, y)^3}{x^3} \int_{\mathfrak{M}} \left(\sum_{n \leq x} e(n\theta) \right)^2 \overline{\left(\sum_{n \leq x} e(n\theta) \right)} d\theta \right| \\ &\ll \Psi(x, y)^3 \frac{\log(u+1)}{\log y} \sum_{q \leq R} \sum_{(a, q)=1} \left(\frac{2^{\omega(q)} q^{1-\alpha} \log^2(q+1)}{\phi(q)} \right)^3 \int_{-R/x}^{R/x} \left(\frac{\log^3(2+|\delta x|)}{(1+|\delta x|)^\alpha} \right)^3 d\delta. \end{aligned}$$

Since $\alpha \geq 0.99$, say, the double sum over q and a is $\ll \sum_{q \leq R} 1/q^{3/2} \ll 1$ and the integral over δ is $\ll \int 1/(1+|\delta x|)^2 \ll 1/x$ (on splitting into the ranges $|\delta| \leq 1/x$ and $|\delta| \geq 1/x$). This gives the desired relation (5.1).

Q.E.D.

As we noted in the Introduction, the error term $O(\log(u+1)/\log y)$ in Corollary 1 only tends to zero if $(\log y)/\log \log x \rightarrow \infty$. To obtain a negligible error term for smaller y (but still satisfying $y \geq \log^K x$), one can replace Major Arc Estimate 2 in our analysis by Proposition 2.3 of Drappeau [12]. There the main term is no longer $(\Psi(x, y)/x)V(x, x; q, \delta)$, but is a more complicated object whose contribution could be analysed as in the conditional work of Lagarias and Soundararajan [28, 29]. We leave this to the interested reader. The point here is that the precise main term for Corollary 1 itself involves the saddle-point $\alpha(x, y)$, and when $1 - \alpha(x, y) = \frac{\log(u \log u) + O(1)}{\log y}$ becomes larger the simple version $\Psi(x, y)^3/2x$ that we stated loses accuracy.

6. PROOF OF COROLLARY 2

To prove Corollary 2 we shall combine Theorem 2, the major arc estimates in §2.2, and the following result, which is essentially proved in §6 of Green's paper [16] but whose precise formulation we take from a slightly later paper of Green and Tao [17].

Transference Principle 1 (See Proposition 5.1 of Green and Tao [17]). *Let N be a large prime, and let $0 < \delta \leq 1$. Let $f : \mathbb{Z}/N\mathbb{Z} \rightarrow [0, \infty)$ and $\nu : \mathbb{Z}/N\mathbb{Z} \rightarrow [0, \infty)$ be any functions such that*

$$f(n) \leq \nu(n) \quad \forall n \in \mathbb{Z}/N\mathbb{Z}, \quad \text{and} \quad \frac{1}{N} \sum_{n \in \mathbb{Z}/N\mathbb{Z}} f(n) \geq \delta.$$

In addition, let $\eta \geq 0$ and $M > 0$ and $2 < p < 3$ be any parameters, and suppose that

$$\left| \frac{1}{N} \sum_{n \in \mathbb{Z}/N\mathbb{Z}} \nu(n) e\left(\frac{an}{N}\right) - \mathbf{1}_{a=0} \right| \leq \eta \quad \forall a \in \mathbb{Z}/N\mathbb{Z}, \quad \text{and} \quad \sum_{a \in \mathbb{Z}/N\mathbb{Z}} \left| \frac{1}{N} \sum_{n \in \mathbb{Z}/N\mathbb{Z}} f(n) e\left(\frac{an}{N}\right) \right|^p \leq M,$$

where $\mathbf{1}$ denotes the indicator function.

Then

$$\frac{1}{N^2} \sum_{n, d \in \mathbb{Z}/N\mathbb{Z}} f(n) f(n+d) f(n+2d) \geq c(\delta) - O_{\delta, M, p}(\eta),$$

where $c(\delta) > 0$ depends on δ only.

Transference Principle 1 asserts that if the non-negative function f has average value at least δ , and the average of the p -th power of $\sum_{n \in \mathbb{Z}/N\mathbb{Z}} f(n) e(an/N)$ is suitably bounded, and if f is *upper bounded* by a function ν whose exponential sums (except the trivial one at $a = 0$) are small enough, then the function f will count quite a lot of three term arithmetic progressions $n, n+d, n+2d$. Note that f and ν are defined on the additive group $\mathbb{Z}/N\mathbb{Z}$, so Transference Principle 1 actually counts arithmetic progressions modulo N . This means that we cannot immediately apply the result to prove Corollary 2, (since the progressions it supplies may wrap around and not be genuine progressions of integers), but there is a standard and easy trick to avoid this problem, as we shall explain below.

For the benefit of an unfamiliar reader, we remark that if $\nu \equiv 1$ then one can take $\eta = 0$, and one can take $p = 2.1$ (say) and $M = 1$ in view of Parseval's identity. Thus Transference Principle 1 readily implies the original form of Roth's theorem, that positive density subsets of the integers contain three term arithmetic progressions. For sparse sets such as the smooth numbers (or the primes) one needs to choose f and ν as functions that grow unboundedly with N , and so verifying the conditions of Transference Principle 1 becomes more challenging.

Recall that in Corollary 2 we are given $B \subseteq \mathcal{S}(y) \cap [1, x]$ such that $\#B \geq \beta\Psi(x, y)$. Choose any large prime $2x < N < 4x$, say, and note that if $n, n+d, n+2d \in B$ modulo N then the representatives $\bar{n}, \overline{n+d}, \overline{n+2d}$ of these numbers in $[1, x]$ must in fact form an arithmetic progression *as integers*, since there is too much space modulo N for any wraparound to occur (indeed we have $\bar{n} + \overline{n+2d} - 2\overline{n+d} \equiv 0$ modulo N , and since $\bar{n} + \overline{n+2d} \in [2, 2x]$ and $-2\overline{n+d} \in [-2x, -2]$ and $N > 2x$, we must in fact have $\bar{n} + \overline{n+2d} - 2\overline{n+d} = 0$ as integers). In particular, if we slightly abuse notation, and let $B : \mathbb{Z}/N\mathbb{Z} \rightarrow \{0, 1\}$ also denote the characteristic function of the set B (modulo N), then to prove Corollary 2 it will suffice to show that

$$\sum_{n \in \mathbb{Z}/N\mathbb{Z}} \sum_{\substack{d \in \mathbb{Z}/N\mathbb{Z}, \\ d \neq 0}} B(n)B(n+d)B(n+2d) \geq 1.$$

We simply set

$$\nu(n) := \frac{N}{\Psi(N, y)} \mathbf{1}_{n \leq N, n \in \mathcal{S}(y)}, \quad \text{and} \quad f(n) := \frac{N}{\Psi(N, y)} B(n),$$

where $\mathbf{1}$ denotes the indicator function. Thus by construction we have $\frac{1}{N} \sum_{n \in \mathbb{Z}/N\mathbb{Z}} \nu(n) = 1$. Moreover, if $a \neq 0$ and if a/N is not within distance $(\log^{20} N)/N$ of a rational with denominator $\leq \log^{20} N$, then Proposition 5 from the previous section yields that

$$\frac{1}{N} \sum_{n \in \mathbb{Z}/N\mathbb{Z}} \nu(n) e\left(\frac{an}{N}\right) = \frac{1}{\Psi(N, y)} \sum_{\substack{n \leq N, \\ n \in \mathcal{S}(y)}} e\left(\frac{an}{N}\right) \ll \frac{1}{\log^5 N} \leq \frac{1}{\log^5 x}.$$

If instead $a \neq 0$ and a/N is within distance $(\log^{20} N)/N$ of a rational with denominator $\leq \log^{20} N$, then Major Arc Estimates 1 and 2 are applicable and can be used to compare $\sum_{\substack{n \leq N, \\ n \in \mathcal{S}(y)}} e\left(\frac{an}{N}\right)$ with $\sum_{n \leq N} e\left(\frac{an}{N}\right)$, implying that

$$\frac{1}{N} \sum_{n \in \mathbb{Z}/N\mathbb{Z}} \nu(n) e\left(\frac{an}{N}\right) = \frac{1}{\Psi(N, y)} \sum_{\substack{n \leq N, \\ n \in \mathcal{S}(y)}} e\left(\frac{an}{N}\right) = \frac{1}{N} \sum_{n \leq N} e\left(\frac{an}{N}\right) + O\left(\frac{\log(u+1)}{\log y}\right) \ll \frac{\log \log x}{\log y},$$

since the final complete sum over n is identically zero.

We obviously always have $f(n) \leq \nu(n)$, and we also note that

$$\frac{1}{N} \sum_{n \in \mathbb{Z}/N\mathbb{Z}} f(n) = \frac{\#B}{\Psi(N, y)} \geq \frac{\beta\Psi(x, y)}{\Psi(N, y)} \geq \frac{\beta\Psi(x, y)}{\Psi(4x, y)} \gg \beta,$$

using Smooth Numbers Result 2. Moreover, we have

$$\begin{aligned} \sum_{a \in \mathbb{Z}/N\mathbb{Z}} \left| \frac{1}{N} \sum_{n \in \mathbb{Z}/N\mathbb{Z}} f(n) e(an/N) \right|^{2.1} &= \frac{1}{\Psi(N, y)^{2.1}} \sum_{a \in \mathbb{Z}/N\mathbb{Z}} \left| \sum_{n \in B} e(an/N) \right|^{2.1} \\ &\ll \frac{N}{\Psi(N, y)^{2.1}} \int_0^1 \left| \sum_{n \in B} e(n\theta) \right|^{2.1} d\theta \\ &\ll 1. \end{aligned}$$

Here the first inequality is a general result of Marcinkiewicz and Zygmund, which the reader may find as Lemma 6.5 of Green [16] (and which is hopefully intuitively plausible, since we might imagine that often $|\sum_{n \in B} e(an/N)|^{2.1} \approx N \int_{a/N-1/2N}^{a/N+1/2N} |\sum_{n \in B} e(n\theta)|^{2.1} d\theta$), and the final inequality follows from Theorem 2.

In summary, we have shown that all the hypotheses of Transference Principle 1 are satisfied, with $\delta = \beta/1000$ (say) and with $p = 2.1$, $M = O(1)$, and $\eta = O((\log \log x)/\log y)$, so we conclude that

$$\begin{aligned} \sum_{n \in \mathbb{Z}/N\mathbb{Z}} \sum_{d \in \mathbb{Z}/N\mathbb{Z}} B(n)B(n+d)B(n+2d) &= \left(\frac{\Psi(N, y)}{N} \right)^3 \sum_{n, d \in \mathbb{Z}/N\mathbb{Z}} f(n)f(n+d)f(n+2d) \\ &\geq \frac{\Psi(N, y)^3}{N} (c(\delta) - O_\delta(\frac{\log \log x}{\log y})). \end{aligned}$$

Provided x and $(\log y)/\log \log x$ are large enough in terms of β the right hand side will be $\gg_\beta \Psi(x, y)^3/x$, whilst the trivial contribution from $d = 0$ to the left hand side is equal to $\#B \leq \Psi(x, y)$. In particular, there will be at least one (and in fact very many) non-trivial three term progressions in B .

Q.E.D.

APPENDIX A. PROOFS OF THE MAJOR ARC ESTIMATES

In this appendix we shall prove the Major Arc Estimates that we stated in §2.2. In all cases, a large portion of the result can be imported immediately from the existing literature (primarily the paper of Drappeau [12]), leaving only a few bad terms or awkward cases to be handled.

A.1. Proof of Major Arc Estimate 1. Under the conditions of Major Arc Estimate 1, if we apply Proposition 2.2 of Drappeau [12] with the choice $Q := \min\{y^{c_2}, e^{c_2\sqrt{\log x}}\}$ we obtain (after suitably relabelling the constant c_2) that

$$\sum_{\substack{n \leq x, \\ n \in \mathcal{S}(y)}} e(n(a/q+\delta)) = V(x, y; q, \delta) + O\left(|W(x, y; q, \delta)| + \Psi(x, y)(1 + |\delta x|)(y^{-c_2} + e^{-2c_2\sqrt{\log x}})\right).$$

We will explain the new term $W(x, y; q, \delta)$ in a moment, but note immediately that for any $A > 0$ we have

$$e^{-2c_2\sqrt{\log x}} = O_A\left(\frac{e^{-c_2\sqrt{\log x}}}{\log^A x}\right) = O_A\left(y^{-c_2} + \frac{e^{-c_2 u/\log^2(u+1)}}{\log^A x}\right),$$

since $u := (\log x)/\log y$ and therefore we always have $e^{-\sqrt{\log x}} \leq \max\{y^{-1}, e^{-u}\}$. This contribution is acceptable for Major Arc Estimate 1, so to prove the estimate it only remains to show that, for any $A > 0$,

$$|W(x, y; q, \delta)| \ll_A \Psi(x, y)(1 + |\delta x|)(y^{-c_2} + \frac{e^{-c_2 u/\log^2(u+1)}}{\log^A x}). \quad (\text{A.1})$$

As explained by Drappeau [12], the term $W(x, y; q, \delta)$ may be omitted unless there exists a real primitive Dirichlet character χ_{bad} , with conductor $r_{\text{bad}} \leq Q = \min\{y^{c_2}, e^{c_2\sqrt{\log x}}\}$, such that r_{bad} divides q and such that the Dirichlet L -function $L(s, \chi_{\text{bad}})$ has a real zero that is $\geq 1 - c/\log Q$ (for a certain absolute constant $c > 0$ which ensures there is at most one such primitive character). If such an exceptional character actually exists, and if we let $\tau(\cdot)$ denote the Gauss sum, and let $\chi_{\text{bad}}^{(d)}$ denote the character to modulus dr_{bad} that is induced by $\chi_{\text{bad}} = \chi_{\text{bad}}^{(1)}$, then $W(x, y; q, \delta)$ is defined by

$$W(x, y; q, \delta) := \frac{\tau(\chi_{\text{bad}})}{\phi(r_{\text{bad}})} \sum_{d|(q/r_{\text{bad}})} \frac{\mu(d)\chi_{\text{bad}}(d)}{\phi(d)} \sum_{\substack{n \leq xr_{\text{bad}}d/q, \\ n \in \mathcal{S}(y)}} e(\delta nq/dr_{\text{bad}}) \chi_{\text{bad}}^{(d)}(n).$$

Using the standard bound $|\tau(\chi_{\text{bad}})| \leq \sqrt{r_{\text{bad}}}$ (as in e.g. chapter 9 of Davenport [10]), we quickly obtain that

$$|W(x, y; q, \delta)| \ll \frac{\sqrt{r_{\text{bad}}} \log x}{\phi(r_{\text{bad}})} \max_{d|(q/r_{\text{bad}})} \left| \sum_{\substack{n \leq xr_{\text{bad}}d/q, \\ n \in \mathcal{S}(y)}} e(\delta nq/dr_{\text{bad}}) \chi_{\text{bad}}^{(d)}(n) \right|.$$

And now we claim that to prove Major Arc Estimate 1, it will suffice to show that

$$\max_{d|(q/r_{\text{bad}})} \left| \sum_{\substack{n \leq xr_{\text{bad}}d/q, \\ n \in \mathcal{S}(y)}} e(\delta nq/dr_{\text{bad}}) \chi_{\text{bad}}^{(d)}(n) \right| \ll \Psi(x, y)(1 + |\delta x|) \log^2 x (y^{-2c_2} + e^{-2c_2 u/\log^2(u+1)}).$$

For if $\log^{c_1} x \leq y \leq e^{\sqrt{\log x}}$, say, then we have $y^{-c_2} \ll 1/\log^3 x$ and $e^{-c_2 u/\log^2(u+1)} \ll_A 1/\log^{A+3} x$, and so our desired bound (A.1) will follow. And if instead $e^{\sqrt{\log x}} < y \leq x$, so that $\log Q = c_2\sqrt{\log x}$, then any exceptional conductor will satisfy $r_{\text{bad}} \gg_A \log^A x$ for any fixed $A > 0$, by Siegel's theorem (as in e.g. chapter 21 of Davenport [10]). Therefore we will get the bound (A.1) because of the prefactor $\frac{\sqrt{r_{\text{bad}}}}{\phi(r_{\text{bad}})} \ll_A 1/\log^A x$.

Next, note that if $q/r_{\text{bad}}d \geq \min\{y^{3c_2}, e^{3c_2u/\log^2(u+1)}\}$ then the trivial bound

$$\left| \sum_{\substack{n \leq xr_{\text{bad}}d/q, \\ n \in \mathcal{S}(y)}} e(\delta nq/dr_{\text{bad}}) \chi_{\text{bad}}^{(d)}(n) \right| \leq \Psi(xr_{\text{bad}}d/q, y) \ll \frac{\Psi(x, y)}{\min\{y^{3c_2}, e^{3c_2u/\log^2(u+1)}\}^{\alpha(x, y)}}$$

is sufficient, (here the second inequality is Smooth Numbers Result 2), since the saddle-point $\alpha(x, y) \geq \alpha(x, \log^{c_1} x) \geq 2/3$ by the approximation (2.1).

If instead $q/r_{\text{bad}}d < \min\{y^{3c_2}, e^{3c_2u/\log^2(u+1)}\}$ then we may split into subsums each of length at most $M := \max\{x/y^{5c_2}, x/e^{5c_2u/\log^2(u+1)}\}$, say, and for any subsum we will have

$$\begin{aligned} \sum_{\substack{x' < n \leq x' + M, \\ n \in \mathcal{S}(y)}} e\left(\frac{\delta nq}{dr_{\text{bad}}}\right) \chi_{\text{bad}}^{(d)}(n) &= e\left(\frac{\delta x'q}{dr_{\text{bad}}}\right) \sum_{\substack{x' < n \leq x' + M, \\ n \in \mathcal{S}(y)}} \chi_{\text{bad}}^{(d)}(n) + O\left(\frac{Mq|\delta|}{dr_{\text{bad}}} \sum_{\substack{x' < n \leq x' + M, \\ n \in \mathcal{S}(y)}} 1\right) \\ &= e\left(\frac{\delta x'q}{dr_{\text{bad}}}\right) \sum_{\substack{x' < n \leq x' + M, \\ n \in \mathcal{S}(y)}} \chi_{\text{bad}}^{(d)}(n) + \\ &\quad + O\left(|\delta x|(y^{-2c_2} + e^{-2c_2u/\log^2(u+1)}) \sum_{\substack{x' < n \leq x' + M, \\ n \in \mathcal{S}(y)}} 1\right). \end{aligned}$$

Here the “big Oh” term is acceptably small.

It only remains to bound the untwisted character sums $\sum_{\substack{x' < n \leq x' + M, \\ n \in \mathcal{S}(y)}} \chi_{\text{bad}}^{(d)}(n)$. However, this has actually already been done in previous work of the author [19]. Indeed, if $\log^{c_1} x \leq y \leq x^{1/(\log \log x)^2}$ (say) then the exceptional characters argument at the end of §3.3 of that paper immediately yields that

$$\begin{aligned} \sum_{\substack{x' < n \leq x' + M, \\ n \in \mathcal{S}(y)}} \chi_{\text{bad}}^{(d)}(n) &= \sum_{\substack{n \leq x' + M, \\ n \in \mathcal{S}(y)}} \chi_{\text{bad}}^{(d)}(n) - \sum_{\substack{n \leq x', \\ n \in \mathcal{S}(y)}} \chi_{\text{bad}}^{(d)}(n) \ll \Psi(x' + M, y) \log^2 x \left(y^{-c} + e^{-cu/\log^2(u+1)}\right) \\ &\ll \Psi(x, y) \log^2 x \left(y^{-c} + e^{-cu/\log^2(u+1)}\right), \end{aligned}$$

for a small absolute constant $c > 0$. This is clearly sufficient, on summing over the $O(\min\{y^{5c_2}, e^{5c_2u/\log^2(u+1)}\})$ values of x' , provided our constant c_2 is small enough in terms of c . If instead $x^{1/(\log \log x)^2} < y \leq x$, then the argument from §3.3 of [19] implies that

$$\sum_{\substack{x' < n \leq x' + M, \\ n \in \mathcal{S}(y)}} \chi_{\text{bad}}^{(d)}(n) \ll \left(\sum_{b|d} \frac{1}{b^\alpha}\right) \Psi(x' + M, y) \frac{\log(r_{\text{bad}})}{\log x} e^{-cu/\log^2(u+1)} \ll \left(\sum_{b|d} \frac{1}{b^\alpha}\right) \Psi(x, y) e^{-cu/\log^2(u+1)}.$$

However, since $d \leq q \leq y^{c_2}$ for some small c_2 , and since $\alpha = \alpha(x, y) = 1 + O(\log(u+1)/\log y)$ by the approximation (2.1), the sum over b is easily seen to be $\ll u \log x \ll \log^2 x$, so again our bound is sufficient.

Q.E.D.

A.2. Proof of Major Arc Estimate 2. Suppose first that $q = 1$, in which case the estimate we are required to prove is that

$$\sum_{\substack{n \leq x, \\ n \in \mathcal{S}(y)}} e(n\delta) = \frac{\Psi(x, y)}{x} \sum_{n \leq x} e(n\delta) + O\left(\Psi(x, y) \frac{\log(u+1)}{\log y} \frac{\log^3(2+|\delta x|)}{(1+|\delta x|)^\alpha}\right), \quad (\text{A.2})$$

whenever $\log^{c_1} x \leq y \leq x$ and $|\delta| \leq \min\{y^{c_2}/x, e^{c_2 \log^{1/4} x}/x\}$. The result is trivial if $\delta = 0$, so suppose throughout that $|\delta| > 0$. Also, let us recall throughout that the saddle-point $\alpha = \alpha(x, y)$ satisfies $\alpha = 1 - O(\frac{\log(u+1)}{\log y})$, by the approximation (2.1), where $u := (\log x)/\log y$.

If $\log^{c_1} x \leq y \leq x^{1/(\log \log x)^4}$ and $|\delta| \leq \min\{y^{c_2}/x, e^{c_2(\log^{3/5} x)/(\log \log x)^{1/5}}/x\}$, for suitable values of c_1, c_2 , then Proposition 2.3 of Drappeau [12] implies that

$$\sum_{\substack{n \leq x, \\ n \in \mathcal{S}(y)}} e(n\delta) = \alpha \Psi(x, y) \int_0^1 e(\delta x t) t^{\alpha-1} dt + O\left(\Psi(x, y) \frac{1}{u} \frac{\log^3(2+|\delta x|)}{(1+|\delta x|)^\alpha}\right).$$

In addition, integration by parts shows that

$$\begin{aligned} \int_0^1 e(\delta x t) t^{\alpha-1} dt &= \left[\frac{e(\delta x t) - 1}{2\pi i \delta x} t^{\alpha-1} \right]_0^1 - (\alpha - 1) \int_0^1 \frac{e(\delta x t) - 1}{2\pi i \delta x} t^{\alpha-2} dt \\ &= \frac{e(\delta x) - 1}{2\pi i \delta x} + O\left((1 - \alpha) \int_0^1 \min\{t, \frac{1}{|\delta x|}\} t^{\alpha-2} dt\right). \end{aligned}$$

Here the integral is clearly $O(1)$ if $|\delta x| \leq 1$, whilst if $1 < |\delta x|$ we split the integral at $t = 1/|\delta x|$ and find it is

$$\ll \frac{1}{|\delta x|^\alpha} + \frac{|\delta x|^{1-\alpha} - 1}{(1 - \alpha)|\delta x|} \ll \frac{\log(2 + |\delta x|)}{(1 + |\delta x|)^\alpha},$$

by distinguishing cases according as $(1 - \alpha) \log(2 + |\delta x|)$ is large or small.

Putting everything together, and also using the bounds $1 - \alpha = O(\frac{\log(u+1)}{\log y})$ and $\left| \frac{e(\delta x) - 1}{2\pi i \delta x} \right| \ll 1/(1 + |\delta x|)$, it follows that

$$\begin{aligned} \sum_{\substack{n \leq x, \\ n \in \mathcal{S}(y)}} e(n\delta) &= \alpha \Psi(x, y) \frac{e(\delta x) - 1}{2\pi i \delta x} + O\left(\Psi(x, y) \left(\frac{1}{u} + \frac{\log(u+1)}{\log y}\right) \frac{\log^3(2+|\delta x|)}{(1+|\delta x|)^\alpha}\right) \\ &= \Psi(x, y) \frac{e(\delta x) - 1}{2\pi i \delta x} + O\left(\Psi(x, y) \left(\frac{1}{u} + \frac{\log(u+1)}{\log y}\right) \frac{\log^3(2+|\delta x|)}{(1+|\delta x|)^\alpha}\right). \end{aligned}$$

By summing the geometric progression we find $\frac{e(\delta x)-1}{2\pi i \delta x} = (1/x) \sum_{n \leq x} e(n\delta) + O(1/x)$. Thus we obtain the claimed estimate (A.2) provided that $\log^{c_1} x \leq y \leq e^{\sqrt{\log x}}$ (say), so that $1/u \ll \log(u+1)/\log y$ in the error term.

We still need to handle the range where $e^{\sqrt{\log x}} < y \leq x$ and $|\delta| \leq \min\{y^{c_2}/x, e^{c_2 \log^{1/4} x}/x\} = e^{c_2 \log^{1/4} x}/x$. To do this we apply Proposition 1 of La Bretèche [5], which implies that

$$\sum_{\substack{n \leq x, \\ n \in \mathcal{S}(y)}} e(n\delta) = \rho(u) \sum_{n \leq x} e(n\delta) + O\left(\Psi(x, y) \frac{\log(u+1) \log(2+|\delta x|)}{\log y (1+|\delta x|)}\right)$$

provided $|\delta| \leq e^{c\sqrt{\log y}}/x$, for a certain small constant $c > 0$. Here $\rho(u)$ is the Dickman function, so as described in §2.1 the main term $\rho(u) \sum_{n \leq x} e(n\delta)$ is indeed

$$\frac{\Psi(x, y)}{x} \left(1 + O\left(\frac{\log(u+1)}{\log y}\right)\right) \sum_{n \leq x} e(n\delta) = \frac{\Psi(x, y)}{x} \sum_{n \leq x} e(n\delta) + O\left(\Psi(x, y) \frac{\log(u+1)}{\log y} \frac{1}{1+|\delta x|}\right),$$

which establishes (A.2). Actually Proposition 1 of La Bretèche [5] is restricted to the range $e^{\log^{2/3+\epsilon} x} \leq y \leq x$, but almost all of the proof works under the much weaker condition that $y \geq e^{(\log \log x)^{5/3+\epsilon}}$. The assumption that $y \geq e^{\log^{2/3+\epsilon} x}$ is only made to ensure that $\Psi(x/e^{c\sqrt{\log y}}, y) \ll \Psi(x, y)/e^{c\sqrt{\log y}}$, and this is certainly also true for all $e^{\sqrt{\log x}} < y \leq x$ (using e.g. Smooth Numbers Result 2 and the approximation (2.1)).

It remains for us to prove Major Arc Estimate 2 in the case where $2 \leq q \leq y^{1/4}$. However, this will be a quick deduction from (A.2). For by definition we see $V(x, y; q, \delta)$ is

$$:= \sum_{\substack{n \leq x, \\ n \in \mathcal{S}(y)}} \frac{\mu(q/(q, n))}{\phi(q/(q, n))} e(n\delta) = \sum_{h|q} \frac{\mu(q/h)}{\phi(q/h)} \sum_{\substack{m \leq x/h, \\ m \in \mathcal{S}(y), \\ (m, q/h)=1}} e(mh\delta) = \sum_{h|q} \frac{\mu(q/h)}{\phi(q/h)} \sum_{d|(q/h)} \mu(d) \sum_{\substack{n \leq x/dh, \\ n \in \mathcal{S}(y)}} e(ndh\delta),$$

where the final equality uses the fact that $\mathbf{1}_{(m, q/h)=1} = \sum_{d|(m, q/h)} \mu(d)$. And we can use our result (A.2) to estimate all the inner sums, deducing that $V(x, y; q, \delta)$ is equal to

$$\sum_{h|q} \frac{\mu(q/h)}{\phi(q/h)} \sum_{d|(q/h)} \mu(d) \left(\frac{\Psi(\frac{x}{dh}, y)}{x/dh} \sum_{n \leq x/dh} e(ndh\delta) + O\left(\Psi\left(\frac{x}{dh}, y\right) \frac{\log(u+1) \log^3(2+|\delta x|)}{\log y (1+|\delta x|)^\alpha}\right) \right).$$

Using Smooth Numbers Result 2 in the form $\Psi(x/dh, y) \ll \Psi(x, y)/(dh)^\alpha \leq \Psi(x, y)q^{1-\alpha}/dh$, we see the contribution from all the “big Oh” error terms is

$$\begin{aligned} &\ll \Psi(x, y) \frac{\log(u+1) \log^3(2+|\delta x|)}{\log y (1+|\delta x|)^\alpha} q^{1-\alpha} \sum_{h|q} \frac{|\mu(q/h)|}{\phi(q/h)h} \sum_{d|(q/h)} \frac{1}{d} \\ &\ll \Psi(x, y) \frac{\log(u+1) \log^3(2+|\delta x|)}{\log y (1+|\delta x|)^\alpha} \frac{2^{\omega(q)} q^{1-\alpha} \log(q+1)}{\phi(q)}, \end{aligned}$$

which is more than acceptable for Major Arc Estimate 2.

Finally, it follows from Théorème 2.4(ii) (last part) of La Bretèche and Tenenbaum [7], and then the approximation $\alpha = 1 - O(\log(u+1)/\log y)$, that $\Psi(x/dh, y)$ is equal to

$$\begin{aligned} (1 + O\left(\frac{\log(u+1)}{\log y} + \frac{\log(dh+1)}{\log x}\right)) \frac{\Psi(x, y)}{(dh)^\alpha} &= \frac{\Psi(x, y)}{dh} (1 + O(\frac{(dh)^{1-\alpha} \log(dh+1) \log(u+1)}{\log y})) \\ &= \frac{\Psi(x, y)}{dh} (1 + O(\frac{q^{1-\alpha} \log(q+1) \log(u+1)}{\log y})). \end{aligned}$$

Remembering that $|\sum_{n \leq x/dh} e(ndh\delta)| \ll x/(dh(1+|\delta x|))$, the contribution from all these “big Oh” terms to $V(x, y; q, \delta)$ may be bounded exactly as we did above. This only leaves the main term contribution, which is

$$\sum_{h|q} \frac{\mu(q/h)}{\phi(q/h)} \sum_{d|(q/h)} \mu(d) \frac{\Psi(x, y)}{x} \sum_{n \leq x/dh} e(ndh\delta) = \frac{\Psi(x, y)}{x} V(x, x; q, \delta),$$

exactly as claimed in Major Arc Estimate 2.

Acknowledgements. The author would like to thank Régis de la Bretèche and Sary Drappeau for their detailed comments and suggestions about a draft of this paper. He also thanks David Conlon for some comments about the literature; Junsoo Ha for explaining his results; Andrew Granville, Ben Green, and K. Soundararajan for their comments and encouragement; and the anonymous referee for his or her comments.

REFERENCES

- [1] A. Balog, C. Pomerance. The distribution of smooth numbers in arithmetic progressions. *Proc. Amer. Math. Soc.*, **115**, no. 1, pp 33-43. 1992
- [2] A. Balog, A. Sárközy. On sums of integers having small prime factors. I. *Studia. Sci. Math. Hungar.*, **19**, no. 1, pp 35-47. 1984
- [3] J. Bourgain. On $\Lambda(p)$ -subsets of squares. *Israel J. Math.*, **67**, no. 3, pp 291-311. 1989
- [4] R. de la Bretèche. Sommes d’exponentielles et entiers sans grand facteur premier. *Proc. London Math. Soc.* (3), **77**, no. 1, pp 39-78. 1998
- [5] R. de la Bretèche. Sommes sans grand facteur premier. *Acta Arith.*, **88**, no. 1, pp 1-14. 1999
- [6] R. de la Bretèche, A. Granville. Densité des friables. *Bull. Soc. Math. France*, **142**, no. 2, pp 303-348. 2014
- [7] R. de la Bretèche, G. Tenenbaum. Propriétés statistiques des entiers friables. *The Ramanujan Journal*, **9**, pp 139-202. 2005
- [8] J. Brüdern. A problem in additive number theory. *Math. Proc. Cambridge Philos. Soc.*, **103**, no. 1, pp 27-33. 1988
- [9] D. Conlon, J. Fox, Y. Zhao. A relative Szemerédi theorem. *Geom. Funct. Anal.*, **25**, no. 3, pp 733-762. 2015
- [10] H. Davenport. *Multiplicative Number Theory*. Third edition, revised by H. L. Montgomery, published by Springer. 2000.
- [11] S. Drappeau. Sur les solutions friables de l’équation $a + b = c$. *Math. Proc. Cambridge Philos. Soc.*, **154**, no. 3, pp 439-463. 2013

- [12] S. Drappeau. Sommes friables d'exponentielles et applications. *Canad. J. Math.*, **67**, no. 3, pp 597-638. 2015
- [13] E. Fouvry, G. Tenenbaum. Entiers sans grand facteur premier en progressions arithmetiques. *Proc. London Math. Soc.* (3), **63**, no. 3, pp 449-494. 1991
- [14] J. B. Friedlander. Integers without large prime factors. II. *Acta Arith.*, **39**, no. 1, pp 53-57. 1981
- [15] A. Granville. Smooth numbers: computational number theory and beyond. *Algorithmic number theory: lattices, number fields, curves and cryptography*, Math. Sci. Res. Inst. Publ., vol. 44, Cambridge University Press, pp 267-323. 2008
- [16] B. Green. Roth's theorem in the primes. *Ann. of Math.* (2), **161**, no. 3, pp 1609-1636. 2005
- [17] B. Green, T. Tao. Restriction theory of the Selberg sieve, with applications. *J. Théor. Nombres Bordeaux*, **18**, no. 1, pp 147-182. 2006
- [18] J. Ha. Some problems in multiplicative number theory. Stanford University PhD thesis, see <http://searchworks.stanford.edu/view/10408988>
- [19] A. J. Harper. Bombieri–Vinogradov and Barban–Davenport–Halberstam type theorems for smooth numbers. Preprint available online at <http://arxiv.org/abs/1208.5992>
- [20] H. A. Helfgott. Minor arcs for Goldbach's problem. Preprint available online at <http://arxiv.org/abs/1205.5252>
- [21] A. Hildebrand. On the number of positive integers $\leq x$ and free of prime factors $> y$. *J. Number Theory*, **22**, no. 3, pp 289-307. 1986
- [22] A. Hildebrand, G. Tenenbaum. On integers free of large prime factors. *Trans. Amer. Math. Soc.*, **296**, no. 1, pp 265-290. 1986
- [23] A. Hildebrand, G. Tenenbaum. Integers without large prime factors. *J. Théor. Nombres Bordeaux*, **5**, no. 2, pp 411-484. 1993
- [24] E. Keil. On a diagonal quadric in dense variables. *Glasg. Math. J.*, **56**, no. 3, pp 601-628. 2014
- [25] E. Keil. Translation invariant quadratic forms in dense sets. Preprint available online at <http://arxiv.org/abs/1308.6680>
- [26] Y. Kohayakawa, T. Łuczak, V. Rödl. Arithmetic progressions of length three in subsets of a random set. *Acta Arith.*, **75**, no. 2, pp 133-163. 1996
- [27] Y. Kohayakawa, V. Rödl, M. Schacht, J. Skokan. On the triangle removal lemma for subgraphs of sparse pseudorandom graphs. *An irregular mind*, Bolyai Soc. Math. Stud., vol. 21, János Bolyai Math. Soc., Budapest, pp 359-404. 2010
- [28] J. C. Lagarias, K. Soundararajan. Smooth solutions to the abc equation: the xyz conjecture. *J. Théor. Nombres Bordeaux*, **23**, no. 1, pp 209-234. 2011
- [29] J. C. Lagarias, K. Soundararajan. Counting smooth solutions to the equation $A+B=C$. *Proc. London Math. Soc.* (3), **104**, no. 4, pp 770-798. 2012
- [30] M. Mirek. Roth's Theorem in the Piatetski–Shapiro primes. *Rev. Mat. Iberoam.*, **31**, no. 2, pp 617-656. 2015
- [31] H. L. Montgomery. *Ten Lectures on the Interface Between Analytic Number Theory and Harmonic Analysis*. Published for the Conference Board of the Mathematical Sciences by the American Mathematical Society. 1994
- [32] G. Sárközy. On sums with small prime factors. *Acta Math. Hungar.*, **67**, no. 4, pp 333-345. 1995

JESUS COLLEGE, CAMBRIDGE CB5 8BL, ENGLAND

E-mail address: A.J.Harper@dpmms.cam.ac.uk